

УДК 004.4

DOI <https://doi.org/10.32782/2521-6643-2026-1-71.9>

**Крамар Ю. М.**, кандидат технічних наук,  
доцент кафедри інформатики та програмної інженерії  
Національного технічного університету України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
ORCID: 0000-0001-7456-726X

**Вітківська І. І.**, старший викладач кафедри  
інформатики та програмної інженерії  
Національного технічного університету України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
ORCID: 0000-0002-9453-9423

**Жаріков Е. В.**, доктор технічних наук,  
професор кафедри інформатики та програмної інженерії  
Національного технічного університету України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
ORCID: 0000-0003-1811-9336

**Радіонов П. Ю.**, кандидат економічних наук,  
доцент кафедри інформатики та програмної інженерії  
Національного технічного університету України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
ORCID: 0000-0002-7164-4458

## МЕТОД ФОРМУВАННЯ ГРУП А/В-ТЕСТІВ ІЗ МІНІМАЛЬНИМ ВЗАЄМНИМ ВПЛИВОМ ІЗ ЗАСТОСУВАННЯМ АЛГОРИТМУ СТОЕРА–ВАГНЕРА

У статті розглядаються сучасні методи формування експериментальних груп для А/В-тестів у цифрових продуктах з урахуванням мінімізації взаємного впливу між учасниками та запобігання spillover-ефектам. Проблема неконтрольованої взаємодії користувачів стає особливо актуальною в умовах масштабних соціальних платформ і висококонективних графів, де поведінка одного індивіда може непрямо впливати на поведінку інших, спотворюючи результати експерименту та знижуючи достовірність висновків. У статті здійснено огляд ключових підходів, що застосовуються у практиці А/В-тестування, включно з кластерною рандомізацією, географічними експериментами та методами стабілізації оцінок у мережесих структурах. Додатково проаналізовано практичні рішення, які використовують Meta та Google для зменшення міжкористувацьких перетинів та покращення чистоти експериментів, зокрема формування стійких кластерів взаємодії та застосування багаторівневого рандомізаційного дизайну.

У роботі запропоновано метод, що ґрунтується на поданні користувачів як графа взаємодії та подальшому формуванні експериментальних груп на основі знаходження мінімального розрізу. Для обчислення оптимального розбиття використано алгоритм Стоера–Вагнера, який забезпечує глобально мінімальний розріз неорієнтованого графа та дозволяє автоматично визначати групи з найменшим можливим міжгруповим впливом. Такий підхід природно відображає реальну структуру соціальних взаємодій і придатний для застосування в системах із мільйонами користувачів. Наведено приклади ваг ребер взаємодій та інтерпретації отриманих кластерів у контексті оцінки spillover-ефектів. Запропонований метод поєднує строгий математичний апарат і практичну релевантність, узгоджуючись із сучасними індустріальними практиками провідних технологічних компаній та забезпечуючи підвищення точності й надійності А/В-експериментів у мережесих середовищах.

Ключові слова: А/В-тест, spillover-ефект, кластеризація, гео-рандомізація, поведінкове групування.

**Kramar Yu. M., Vitkovska I. I., Zharikov E. V., Radionov P. Yu. Method for forming A/B test groups with minimal interference using the Stoer–Wagner algorithm**

*This article examines contemporary methods for constructing experimental groups in A/B testing for digital products, with a specific focus on minimizing mutual interference between participants and reducing spillover effects arising from social or behavioral interactions. The challenge of uncontrolled user-to-user influence has grown increasingly important in highly connected online environments, where actions, recommendations, and shared content can propagate across a network and distort*



© Ю. М. Крамар, І. І. Вітківська, Е. В. Жаріков, П. Ю. Радіонов, 2026  
Стаття поширюється на умовах ліцензії відкритого доступу CC BY 4.0

---

*the measured impact of an experimental treatment. Such interference undermines the validity of causal inference and often leads to biased or unstable conclusions about product performance.*

*The paper provides an overview of established approaches, including cluster-based randomization, geographic experiments, and hierarchical assignment strategies, as well as practical implementations widely adopted by leading technology companies such as Meta and Google. These organizations regularly face large-scale network effects and have developed robust methodologies to ensure reliable experimentation even when user interactions cannot be ignored.*

*A central contribution of the article is the introduction of a graph-based method for constructing experimental groups. In this approach, users are represented as nodes within an interaction graph, and the strength of their connections reflects the probability or intensity of potential interference. Group formation is achieved through identifying a minimum cut in this graph, computed using the Stoer–Wagner algorithm. This algorithm enables efficient and scalable partitioning of large networks while guaranteeing an optimally minimal sum of cross-group connections. As a result, the proposed method effectively reduces spillover effects and significantly improves the internal validity of A/B tests.*

*The advantages of minimum-cut partitioning include natural alignment with real-world interaction patterns, suitability for graphs containing millions of users, reduced risk of contamination between treatment and control groups, and consistency with experimentation frameworks observed in industry practice. These benefits underscore the relevance and applicability of graph-based experimental design for modern digital ecosystems.*

*Key words: A/B testing, spillover effect, clustering, geo-randomization, behavioral grouping.*

**Постановка проблеми.** А/В-тестування – один з найважливіших інструментів вимірювання ефективності змін у цифрових продуктах. Його класична модель передбачає, що користувачі незалежні між собою, а учасники експериментальних груп не впливають одне на одного. Однак сучасні інтернет-платформи є мережевими: користувачі взаємодіють, коментують, діляться контентом, змагаються за спільні ресурси або рекламу. У таких умовах незалежність порушується, а результати А/В-тесту можуть спотворюватися. Це робить звичайні методи випадкового розбиття недостатніми.

Це спотворення називається spillover-ефектом – ситуацією, коли зміна, застосована до однієї групи користувачів, побічно впливає на поведінку користувачів іншої групи [1, 2]. Наприклад, якщо користувач з групи А поширює контент у соціальній мережі, то друзі з групи В можуть побачити наслідки експерименту, хоча вони не є його учасниками.

Соціальні мережі (Meta), рекомендаційні системи (YouTube), рекламні ринки (Google Ads) та інтерактивні платформи мають сильну мережеву природу. У таких випадках spillover-ефекти – це дуже розповсюджена і звичайна річ. Саме тому сучасні технологічні компанії (Meta, Google, Amazon, TikTok) активно розвивають та застосовують методи, які дозволяють формувати ізольовані експериментальні групи, мінімізуючи взаємний вплив [3–9]. Ці методи варіюють від гео-рандомізації до складних алгоритмів розбиття графів.

Отже, у статті розглядаються загальні підходи та їхня практична реалізація, а також пропонується один із методів – формування груп через пошук мінімального розрізу на графі взаємодій за допомогою алгоритму Стоера–Вагнера [10].

#### **Методи мінімізації взаємного впливу між групами А/В-тестів**

Розглянемо поширені методи мінімізації взаємного впливу між групами А/В-тестів та випадки їх використання.

##### **1. Географічна рандомізація (Geo-experiments)**

Метод базується на розбитті користувачів не індивідуально, а агреговано – за регіонами, домогосподарствами або DMA-зонами. Цей метод використовується компанією Google у власних продуктах: Google Ads (оцінка змін в аукціонах), YouTube (регіональні зміни рекомендацій), Google Maps (тестування локальних функцій) [3, 4].

Перевагою методу є мінімальні перетини між геокластерами, що призводить до зменшення spillover-ефекту.

##### **2. Кластеризація соціального графа**

Компанія Meta широко застосовує розбиття користувачів на спільноти за допомогою алгоритмів Louvain, Leiden, METIS [6, 7]. В основі цих алгоритмів лежить ідея, що кластер являє собою групу користувачів, які щільно взаємодіють, а експериментальні групи формуються за принципом «кластер-до-групи», а не «користувач-до-групи».

Метод дає змогу уникнути змішування тісно пов'язаних користувачів, зменшити вплив поведінки друзів на результати тесту, коректно вимірювати ефекти у стрічці новин (Feed).

##### **3. Двоступенева мережева рандомізація (Two-stage randomization)**

Компанія Meta застосовує цей метод у складних мережевих експериментах: на першому етапі обираються кластери/спільноти, що братимуть участь у тесті; на другому – усередині частини кластерів рандомізують користувачів [8, 9].

Метод дозволяє вимірювати прямі ефекти (на користувача), непрямі ефекти (через друзів)

##### **4. Розбиття на основі поведінкових або структурних ознак**

Компанії Google і Meta застосовують наступні моделі: спільні пристрої (household graph), переглядові подібності (view graph у YouTube), модельовані вектори поведінки (embeddings) [3, 4, 8, 9].

---

Ціль розбиття: користувачі з високою кореляцією поведінки повинні або бути в одній групі, або не мати перетинів.

#### **Практичне застосування методів мінімізації взаємного впливу**

Як приклад, компанія Meta виконала ряд експериментів на соціальному графі [5]. При тестуванні алгоритму ранжування стрічки потрібно було контролювати мережевий вплив. Для цього побудовано граф взаємодій користувачів, виконано його кластеризацію, проведено експерименти на рівні кластерів.

Такий підхід дозволив уникнути ситуації, коли контент із експериментальної групи з'являється у стрічці користувача з контрольною.

Ще один приклад від компанії Google, що використала гео-тести та графи рекламодавців [3, 4]. У Google Ads рекламодавці конкурують в одних аукціонах і їх неможливо незалежно рандомізувати. Тому застосовуються geo-split тестування та графове розбиття рекламодавців за допомогою METIS [7].

Такий підхід дозволив мінімізувати перетин між рекламодавцями у групах.

Всі наведені у статті методи – кластеризація, гео-рандомізація, поведінкове групування – мають на меті мінімізувати кількість або силу зв'язків між експериментальними групами і застосовуються досить успішно. Але найбільш формалізованим, математично обґрунтованим способом реалізувати цю мету є побудова графу взаємодій користувачів та знаходження на ньому глобального мінімізованого розрізу між двома підмножинами.

Графовий метод мінімального розрізу є ефективним у випадку, коли: система мережево структурована; є вимірні взаємодії між користувачами; spillover-ефекти значні та вимірювані.

Саме цей підхід добре масштабується на соціальні графи, графи спільних переглядів, графи пристроїв, графи конкуренції рекламодавців і забезпечує формально мінімальний spillover – ефект.

#### **Метод формування груп на основі графа взаємодій**

Метод базується на представленні користувачів як графа взаємодій для формування груп через мінімальний розріз.

##### **1. Побудова графа.**

Нехай є граф

$$G = (V, E, w), \quad (1)$$

де  $V$  – множина користувачів;  $E$  – множина ребер (факт взаємодії);  $w(u, v)$  – вага ребра, що відображає силу впливу.

Розглянемо приклад ваги ребра  $w(u, v)$ , в якому між користувачами  $u$  і  $v$  існують такі взаємодії за останні 7 днів:

- користувачі написали один одному 5 повідомлень (вплив взаємодії визначимо, наприклад, як 0.6);
- 3 рази відмітили (поставили «like») один одному пости (вплив – 0.3);
- 2 рази переглядали профіль один одного (вплив – 0.1);
- 1 раз поширили спільний контент.

Тоді вагу можна визначити як:

$$w(u, v) = 0.6 \cdot \text{messages}(u, v) + 0.3 \cdot \text{likes}(u, v) + 0.1 \cdot \text{profile\_views}(u, v). \quad (2)$$

Підставимо в формулу значення:

$$w(u, v) = 0.6 \cdot 5 + 0.3 \cdot 3 + 0.1 \cdot 2 = 3 + 0.9 + 0.2 = 4.1.$$

Таким чином, ребро  $(u, v)$  матиме вагу 4.1, що означає відносно сильний взаємний вплив, тому небажано розділяти користувачів  $u$  і  $v$  між експериментальними групами.

##### **2. Пошук мінімального розрізу**

Потрібно знайти такі множини  $S$  і  $T$ , що:

$$\text{cut}(S, T) = \sum_{\{u \in S, v \in T\}} w(u, v). \quad (3)$$

Мінімальний розріз надає мінімальний взаємний вплив. Для вирішення задачі пошуку мінімального розрізу в даній статті пропонується використати алгоритм Стоера–Вагнера [6].

#### **Знаходження глобального мінімізованого розрізу алгоритмом Стоера–Вагнера**

Алгоритм Стоера–Вагнера (Stoer–Wagner Minimum Cut Algorithm) – це детермінований алгоритм для знаходження глобального мінімального розрізу у зваженому ненаправленому графі. Алгоритм забезпечує оптимальний результат та є одним із ефективних вирішень для задач такого типу [10].

Мінімальний розріз – це спосіб розділити граф на дві частини так, щоб сумарна вага «розірваних» ребер була мінімальною.

Алгоритм працює методом багаторазового «стягування» графа, відповідно до якого:

- виконується пошук пари вершин, між якими найбільший зв'язок, збільшуючи набір вершин ( $A$ ) у певному порядку;

- останні дві додані вершини вважаються кандидатним мінімальним розрізом;
- вище вказані дві вершини зливаються, утворюючи одну нову вершину;

- процес повторюється, доки в графі не залишиться одна вершина;
- кращий розріз серед усіх ітерацій – і є глобальний мінімальний розріз.

Розглянемо формальне визначення алгоритму Стоера–Вагнера.

Нехай є граф (описаний вище),  $G = (V, E, w)$ , де  $w(u, v)$  – вага ребра між вершинами  $(u)$  і  $(v)$ .

Алгоритм складається з  $n - 1$  фаз (ітерацій) (де  $n$  – це кількість вершин).

Фазу алгоритму представлено послідовністю дій:

- задаємо початкову множину  $A$ :  $A = \{\text{довільна вершина}\}$ ;
- поступово додаємо найбільш зв'язані вершини. На кожному кроці додається вершина, яка має найбільшу сумарну вагу зв'язків до поточного набору  $(A)$ ;
- знаходимо останні дві додані вершини, та позначимо їх як  $(s)$  і  $(t)$  (останньою додається  $(t)$ ), тоді множина, що містить тільки  $(t)$ , описує кандидатний розріз

$$cut_t = \sum_{u \in A'} w(u, t), \quad (4)$$

де  $(A')$  – всі вершини, додані до  $(A)$  до  $(t)$ ;

- оновлюємо найкращий мінімальний розріз. Якщо  $cut_t$  найменший з усіх фаз, то зберігаємо його;
- зливаємо вершини  $(s)$  і  $(t)$ . Це зменшує кількість вершин у графі та створює новий агрегований вузол з вагами:

$$w'(u, st) = w(u, s) + w(u, t);$$

- переходимо до наступної фази (тобто робимо ітерацію всіх кроків), поки не залишиться 1 вершина.

В результаті виконання алгоритму отримуємо:

- множини вершин  $S$  і  $T$ , які задають мінімальний розріз,
- значення ваги цього розрізу, що є сумою перетнутих ребер.

Особливості алгоритму Стоера–Вагнера:

- алгоритм виявляє «щільно зв'язані групи» через поступове додавання найсильніших зв'язків;
- у момент, коли додається остання вершина  $(t)$ , вона має найменший зв'язок із зовнішніми вершинами, тому розріз, що відокремлює її, є кандидатом на мінімальний;
- багаторазове злиття вершин не порушує структуру мінімального розрізу, оскільки агрегована вершина зберігає сумарну силу зв'язків;
- перебір проводиться у всіх фазах, що гарантує знаходження глобального мінімуму.

Перевагами методу мінімального розрізу є:

- мінімальний можливий міжгруповий вплив;
- природне відображення реальної структури взаємодій;
- здатність застосовуватись до великих графів (мільйони користувачів);
- узгодженість з практиками Meta та Google.

**Висновки.** У сучасних цифрових системах взаємний вплив між користувачами є ключовою проблемою А/В-тестування. Існуючі практики Meta та Google – гео-рандомізація, кластеризація соціальних графів, двоступенева мережева рандомізація – демонструють потребу у методах, що враховують мережеву природу продуктів.

Найбільш універсальним і математично строгим підходом є представлення користувачів як графа взаємодій та формування груп через мінімальний розріз. Алгоритм Стоера–Вагнера дозволяє знаходити оптимальне розбиття, що мінімізує spillover-ефекти та забезпечує високу якість експерименту.

Інтеграція цього методу у практику А/В-тестування створює потужний інструмент для побудови ізольованих, незалежних і статистично валідних експериментальних груп.

#### Список використаних джерел:

1. Xu Y. From Infrastructure to Culture: A/B Testing Challenges in Large Scale Social Networks / Xu Y., Nanyu Ch., Fernandez A., Sinno O., Bhasin A. 2015. URL: <https://dl.acm.org/doi/abs/10.1145/2783258.2788602>
2. Holmström Olsson H. Experimentation that Matters: A Multi-case Study on the Challenges with A/B Testing. Holmström Olsson H., Bosch J., Fabijan A. 2017. URL: [https://link.springer.com/chapter/10.1007/978-3-319-69191-6\\_12](https://link.springer.com/chapter/10.1007/978-3-319-69191-6_12)
3. Vaver J. Measuring Ad Effectiveness Using Geo Experiments. Vaver J., Koehler J. Google. URL: <https://static.googleusercontent.com/media/research.google.com/en/pubs/archive/38355.pdf>
4. Kerman J. Estimating Ad Effectiveness using Geo Experiments in a Time-Based Regression Framework. Kerman J., Wang P., Vaver J. Google. URL: <https://static.googleusercontent.com/media/research.google.com/en/pubs/archive/45950.pdf>
5. Karrer B. Network experimentation at scale. Karrer B., Shi L., Bhole M., Goldman M., Palmer T., Gelman C., Konutgan M., Sun F. Cornell University, 2020. URL: <https://arxiv.org/abs/2012.08591>

- 
6. Karypis G. A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs. Karypis G., Kumar V. *SIAM Journal on Scientific Computing*. 1998. Т. 20, № 1. С. 359–392. <https://doi.org/10.1137/S1064827595287997>
  7. Blondel V. D. Fast Unfolding of Communities in Large Networks. Blondel, V. D., & Guillaume, J. L. *Journal of Statistical Mechanics*. 2008. Режим <https://doi.org/10.1088/1742-5468/2008/10/P10008>
  8. How Meta tests products with strong network effects. URL: <https://medium.com/%40AnalyticsAtMeta/how-meta-tests-products-with-strong-network-effects-96003a056c2c>.
  9. Testing Product Changes with Network Effects. URL: <https://research.facebook.com/blog/2021/8/testing-product-changes-with-network-effects>
  10. Stoer M. A simple min-cut algorithm. / Stoer M., Wagner F. *Journal of the ACM*. 1997. Т. 44, № 4. С. 585–591. <https://doi.org/10.1145/263867.263872>

#### References:

1. Xu, Y., Nanyu, C., Fernandez, A., Sinno, O., & Bhasin, A. (2015). From infrastructure to culture: A/B testing challenges in large scale social networks. *KDD '15: The 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Retrieved from: <https://dl.acm.org/doi/abs/10.1145/2783258.2788602>
2. Holmström Olsson, H., Bosch, J., & Fabijan, A. (2017). Experimentation that matters: A multi-case study on the challenges with A/B testing. *Software Business* (pp. 179–185). Retrieved from: [https://doi.org/10.1007/978-3-319-69191-6\\_12](https://doi.org/10.1007/978-3-319-69191-6_12)
3. Vaver, J., & Koehler, J. (2011). Measuring ad effectiveness using geo experiments. *Google Research*. Retrieved from: <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/38355.pdf>
4. Kerman, J., Wang, P., & Vaver, J. (2017). Estimating ad effectiveness using geo experiments in a time-based regression framework. *Google Research*. Retrieved from: <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45950.pdf>
5. Karrer, B., Shi, L., Bhole, M., Goldman, M., Palmer, T., Gelman, C., Konutgan, M., & Sun, F. (2020). Network experimentation at scale. *Cornell University arXiv*. Retrieved from: <https://arxiv.org/abs/2012.08591>
6. Karypis, G., & Kumar, V. (1998). A fast and high-quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, 20(1), 359–392. <https://doi.org/10.1137/S1064827595287997>
7. Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
8. Analytics at Meta. (2022). How Meta tests products with strong network effects. *Medium*. Retrieved from: <https://medium.com/@AnalyticsAtMeta/how-meta-tests-products-with-strong-network-effects-96003a056c2c>
9. Karrer B., Shi L., Bhole M. (2021). Testing product changes with network effects. *Meta Research*. Retrieved from: <https://research.facebook.com/blog/2021/8/testing-product-changes-with-network-effects/>
10. Stoer, M., & Wagner, F. (1997). A simple min-cut algorithm. *Journal of the ACM*, 44(4), 585–591. <https://doi.org/10.1145/263867.263872>

Дата першого надходження статті до видання: 23.11.2025

Дата прийняття статті до друку після рецензування: 22.12.2025

Дата публікації (оприлюднення) статті 27.01.2026