

DOI: 10.20535/kpissn.2025.3.339188

УДК 519.8, 004.9

М.О. Щоголев^{1*}, В.В. Циганок^{1, 2}¹КПІ ім. Ігоря Сікорського, Київ, Україна²Інститут проблем реєстрації інформації НАН України, Київ, Україна

*Відповідальний автор: shchoholiev.maksym@gmail.com

МЕТОД КОРИГУВАННЯ ІНДЕКСУ СОЦІАЛЬНОЇ НАПРУЖЕНОСТІ НА ОСНОВІ МЕТОДІВ ВИЗНАЧЕННЯ УЗГОДЖЕНОСТІ ЕКСПЕРТНИХ ОЦІНОК

Проблематика. Соціальна напруженість є предметом багатьох сучасних соціологічних досліджень, оскільки вона має значний вплив на успішність реалізації прийнятих рішень як на державному рівні, так і в бізнесі. Соціальну напруженість можна визначати на основі аналізу реакцій людей на новинні події у соціальних мережах. Отримання узагальнених оцінок впливу новинних подій на підвищення рівня соціальної напруженості ускладнюється тим, що відношення читачів різних інформаційних каналів до однієї й тієї самої новини можуть суттєво різнитись. Проблема застосовності методів визначення узгодженості експертних оцінок до задач, що не використовують експертну інформацію, залишається малодослідженою.

Мета дослідження. Розроблення і тестування методу коригування індексу соціальної напруженості на основі методів визначення узгодженості експертних оцінок.

Методика реалізації. Методика реалізації методу коригування індексу соціальної напруженості полягає у взаємодії бази знань системи підтримки прийняття рішень і методу визначення узгодженості експертних оцінок на основі мінімізації суми відстаней між попарно взятими оцінками.

Результати дослідження. Запропоновано метод коригування індексу соціальної напруженості на основі методів визначення узгодженості експертних оцінок. Запропоновано методику визначення порогового значення узгодженості для оцінок впливу новинних публікацій на підвищення соціальної напруженості. У результаті застосування методу коригування індексу соціальної напруженості достовірність результатів оцінювання соціальної напруженості може бути підвищена.

Висновки. Застосування методу коригування індексу соціальної напруженості на основі методів визначення узгодженості експертних оцінок дало можливість підвищити достовірність визначення соціальної напруженості у випадках, коли індекс узгодженості між оцінками впливів новинних подій нижчий за порогове значення узгодженості.

Ключові слова: соціальна напруженість; соціальні мережі; новинна подія; система підтримки прийняття рішень; узгодженість оцінок; достатній для агрегації рівень узгодженості.

Вступ

Соціальна напруженість (СН) є предметом багатьох сучасних соціологічних досліджень, оскільки вона має значний вплив на успішність реалізації прийнятих рішень як на державному рівні, так і в бізнесі. Причиною СН є наявність невірних соціальних протиріч, що проявляються у формі конфлікту думок, інтересів чи цінностей всередині суспільства. Різке зростання рівня СН спостерігається як реакція суспільства на деякі

новинні події, що стосуються предмету протиріччя і порушують встановлений до того баланс.

Щоб визначити динаміку зміни СН, зумовленої новинними подіями, а також ступінь впливу кожної новини на підвищення її рівня, розроблено математичну модель СН, побудовану на аналізі емоційності коментарів із соціальних мереж. Для побудови математичної моделі використано систему підтримки прийняття рішень мережевої структури, яка розраховує індекс СН у кожний момент часу.

Пропозиція для цитування цієї статті: М.О. Щоголев, В.В. Циганок, «Метод коригування індексу соціальної напруженості на основі методів визначення узгодженості експертних оцінок», Наукові вісті КПІ, № 3, с. 32–41, 2025. doi: <https://doi.org/10.20535/kpissn.2025.3.339188>

Offer a citation for this article: Maksym Shchoholiev, Vitaliy Tsyganok, “A method for adjusting the social tension index based on methods for assessing the consistency of expert estimates”, KPI Science News, no. 3, pp. 32–41, 2025. doi: <https://doi.org/10.20535/kpissn.2025.3.339188>

Однією із проблем, яка виникає у процесі розрахунку цього індексу, є узгоджуваність оцінок, які отримують у результаті аналізу коментарів з різних інформаційних ресурсів (ІР). Причиною виникнення цієї проблеми є відмінності у поглядах читачів різних ІР щодо одних і тих самих новинних подій. Підрахунок узагальненого індексу напруженості на основі оцінок впливів новинних публікацій з низьким ступенем узгодженості може призвести до втрати інформативності цього індексу. Отже, постала проблема розробити методику визначення порогового значення узгодженості для оцінок впливу новинних публікацій на підвищення СН. Порогове значення потрібне для фільтрації тих оцінок впливу новинних публікацій, рівень узгодженості яких досить низький. Низький рівень узгодженості вказує на полярність думок читачів різних ІР, що може потенційно спричинити зростання рівня напруженості, тому наступним постало питання розробити метод коригування індексу СН на основі оцінок впливу новинних публікацій з низьким ступенем узгодженості.

1. Метод визначення соціальної напруженості на основі даних із соціальних мереж

Основними методами визначення рівня СН, які найчастіше використовують сьогодні, є метод соціологічних опитувань і метод статистичних досліджень [1].

Перевагами методу соціологічних опитувань є можливість ставити питання, які прямо стосуються СН, можливість додатково отримати інформацію про респондентів, таку як вік, стать, місце проживання тощо. Найголовнішим недоліком цього методу є те, що результати соціологічних опитувань мають ретроспективний характер. Оскільки рівень СН може швидко змінюватись, а саме опитування та аналіз результатів анкетування потребують певного часу, остаточні результати на момент їх отримання можуть бути вже неактуальними. Важливим недоліком цього методу також є низька часова дискретність, що унеможливорює відстеження швидких коливань рівня напруженості.

Метод статистичних досліджень, що використовує дані переважно державної статистики про рівень безробіття, смертності населення, злочинності тощо, вирізняється тим, що більшість даних для нього є у вільному доступі, і він дозволяє підраховувати індекси напруженості в окремих регіонах країни. Утім, він має низьку суттєвих недоліків, серед яких найважливі-

шими є непрямий характер статистики, тобто дослідження лише факторів, які можуть вплинути на рівень напруженості, а не самі настрої у суспільстві, а також великий часовий розрив між досліджуваним часовим інтервалом і часом публікації аналітичних звітів.

Метод визначення СН на основі даних із соціальних мереж потребує значно більшої технічної підтримки, ніж опитувальні та статистичні методи, але він дозволяє подолати їхні найсуттєвіші недоліки. По-перше, якщо налагоджено механізм збору даних із соціальних мереж в реальному часі, цей метод дозволяє отримувати актуальну інформацію про рівень напруженості в суспільстві з порівняно невеликою затримкою. Завдяки цьому з'являється можливість досліджувати зміну рівня напруженості в динаміці. По-друге, використання даних із соціальних мереж дозволяє автоматизувати процес періодичного визначення рівня СН.

У цій роботі застосовано метод визначення рівня СН, описаний у [2, 3]. Він складається із двох етапів: сентимент-аналізу коментарів із соціальних мереж і підрахунку рівня напруженості в суспільстві на основі оцінок цього сентимент-аналізу.

На першому етапі відбувається збирання даних із соціальних мереж. Вони включають новинні публікації, коментарі до них і метадані, такі як час, інформація про реакції на коментарі, назва ІР тощо. Після цього за допомогою нейронної мережі визначають оцінку емоційності кожного коментаря на інтервалі $[0, 2]$, де 0 – максимум позитивної емоційності, 2 – максимум негативної емоційності. На основі отриманих даних для кожної новинної публікації будують розподіл її коментарів за рівнем емоційності. Ці розподіли для кожної новинної публікації подають на вхід системи підтримки прийняття рішень.

2. Архітектура бази знань системи підтримки прийняття рішень

Система підтримки прийняття рішень (СППР) призначена для визначення ступеня впливу новинних публікацій і новинних подій на підвищення рівня СН, а також для підрахунку загального індексу СН для певного ІР.

База знань розробленої СППР має трирівневу мережеву структуру (рис. 1).

Для кожного ІР будують окрему базу знань.

Усі вузли від 1 до n і від 6 до k зображують критерії СН. Перший (найнижчий) рівень вуз-

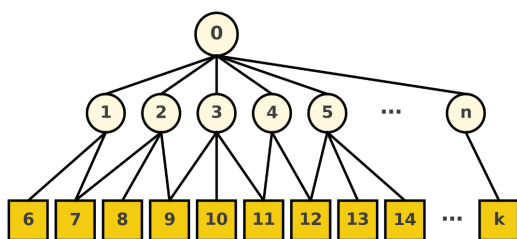


Рис. 1 – Архітектура бази знань СППР

лів, позначений числами від 6 до k , зображує новинні публікації. Другий рівень вузлів, позначений числами від 1 до n , зображує новинні події. Кількість вузлів n і k для різних ІР може бути різною. Вузол з числом 0 зображує головну ціль – підвищення СН.

3. Алгоритм розрахунку індексів соціальної напруженості

Щоб визначити рівень напруженості, розраховують ступінь досягнення головної цілі для кожного ІР. Розрахунок проводять за таким алгоритмом.

1. На основі розподілу оцінок емоційності коментарів для кожної новинної публікації розраховують частковий коефіцієнт впливу (ЧКВ) за формулою:

$$w_j = \sum_{k=1}^n \frac{k}{n} N_k, \quad (1)$$

де w_j – ЧКВ j -ї новинної публікації на новинну подію або події, з якими вона пов'язана; n – кількість рівномірно розподілених оцінок на інтервалі $[0, 2]$; N_k – кількість коментарів, що мають відповідну оцінку емоційності.

2. Частковий коефіцієнт впливу цілей першого рівня (новинних публікацій) нормалізують у групах, що пов'язані з кожною новинною подією.

3. Підраховують ступені досягнення цілей другого рівня (підвищення напруженості, зумовлене новинними подіями) як суми добутків нормалізованих ЧКВ новинних публікацій на коефіцієнти активності користувачів досліджуваної соціальної мережі, які визначають залежно від часу проведення дослідження напруженості:

$$\theta_e = \sum_{j=1}^p w_{e,j} A(t), \quad (2)$$

де θ_e – ступінь досягнення цілі для кожної новинної події e ; $w_{e,j}$ – частковий коефіцієнт впливу j -ї новинної публікації на e -ту новинну подію;

$A(t)$ – коефіцієнт активності користувачів соціальної мережі залежно від часу проведення дослідження; p – кількість публікацій, пов'язаних із e -ю новинною подією.

4. Підраховують ЧКВ цілей другого рівня за формулою, подібною до формули (1):

$$w_e = \sum_{k=1}^n \frac{k}{n} N_k, \quad (3)$$

де w_e – частковий коефіцієнт впливу e -ї новинної події на підвищення рівня напруженості; N_k – загальна кількість коментарів до новинних публікацій, пов'язаних із новинною подією e , що мають оцінку емоційності з номером k .

5. Проводять нормалізацію всіх ЧКВ цілей другого рівня.

6. Підраховують ступінь досягнення головної цілі як суму добутків нормалізованих ЧКВ цілей другого рівня на ступені досягнення цілей другого рівня, тобто

$$\Theta = \sum_{e=1}^E w_e \theta_e = \sum_{e=1}^E v_e, \quad (4)$$

де Θ – ступінь досягнення головної цілі; E – кількість новинних подій, що розглядаються, $w_e \theta_e = v_e$ – внесок події у підвищення СН у межах ІР.

Щоб перейти з рівня одного ІР на рівень загальних оцінок за всіма ІР, треба провести агрегацію значень v_e для кожної цілі другого рівня (новинні події) за всіма ІР, де відповідна подія висвітлювалась у публікаціях. Ступені досягнення цілей другого рівня інтерпретуються як оцінки впливу новинних подій на підвищення рівня напруженості. Під час агрегації враховується вага кожного ресурсу. Будемо вважати, що чим більшу кількість підписників має ресурс, тим більший його вплив на підвищення рівня напруженості. Нехай γ_i – нормалізована до одиничної суми кількість підписників на ІР i . Тоді загальний внесок події у підвищення СН можна підраховувати за формулою

$$\tilde{v}_e = \sum_i v_{e,i} \gamma_i, \quad (5)$$

де $\tilde{v}_e$ – агрегований внесок події e у підвищення рівня СН; $v_{e,i}$ – внесок події e у межах i -го ресурсу.

Тоді для підрахунку загального ступеня досягнення головної цілі за всіма ІР у формулі (4) замість v_e використовують $\tilde{v}_e$. Цей ступінь досягнення головної цілі інтерпретують як індекс СН у момент часу t .

Оскільки відношення читачів різних ІР до однієї й тієї самої новини можуть суттєво різ-

нитись, рівень узгодженості цих оцінок впливу може бути низьким, що може призвести до зниження достовірності агрегованих оцінок для новинних подій. Рівень узгодженості визначає, чи є доцільним проведення агрегації оцінок впливу.

4. Застосування методу визначення узгодженості експертних оцінок

4.1. Аналіз відомих методів

У роботі [4] проведено порівняння наявних індексів узгодженості експертних оцінок. Автори публікації виділяють такі індекси:

- 1) стандартне відхилення вибірки [5];
- 2) коефіцієнт варіації;
- 3) індекс скоригованого середнього відхилення [6];
- 4) індекси r_{wg} [7] та r_{wg}^* [8];
- 5) індекс a_{wg} [9];
- 6) спектральний показник узгодженості [10];
- 7) зважений індекс спарювання [11];
- 8) критерій узгодженості Пірсона;
- 9) індекс подвійної ентропії [12].

Досліджуючи переваги й недоліки цих індексів, автори доходять висновку, що жодний із них не задовольняє базові вимоги, виконання яких може підвищити достовірність експертиз. До цих вимог входять вимоги про:

- 1) існування «істинної» оцінки під час експертного оцінювання;
- 2) відповідність цієї «істинної» оцінки деякому усередненому значенню множини індивідуальних оцінок експертів;
- 3) можливість подати множину індивідуальних оцінок на обмеженій з обох боків числовій шкалі;
- 4) досяжність максимального рівня узгодженості (1.0) тоді і тільки тоді, коли всі експерти вибрали одну й ту ж саму оцінку;
- 5) незалежність індексу узгодженості від зсувів оцінок;
- 6) про те, що збільшення розміру шкали має призводити до підвищення індексу узгодженості за однієї і тієї самої множини оцінок, і навпаки;
- 7) незалежність від розміру шкали: якщо деяка множина оцінок більш узгоджена, ніж інша у певній шкалі, то вона має бути більш узгодженою і в будь-якій іншій шкалі;
- 8) масштабованість індексу: у разі лінійних змін (одночасного пропорційного збільшення/зменшення) розмірів шкали і значень усіх оцінок індекс узгодженості має залишатись незмінним.

Усі ці вимоги є так само актуальними і для індексів узгодженості, що будуть використані для оцінок впливу новинних подій.

Автори у [4] пропонують власний метод підрахунку індексів узгодженості, що задовольняють усі висунуті вимоги. Індеси підраховують за формулою

$$I = 1 - \frac{\sum_{i \neq j} f(|x_i - x_j|)}{M}, \quad (6)$$

де x_i — оцінка i -го експерта; M — значення для найбільш неузгодженого випадку, тобто таке, що максимізує $\sum_{i \neq j} f(|x_i - x_j|)$, $I \in [0, 1]$. Для функції f мають виконуватись такі вимоги: $f(x) > 0$, якщо $x > 0$, та $f(x) = 0$, якщо $x = 0$. Ці вимоги виконуються, наприклад, для функцій $f(x) = x^2$ та $f(x) = |x|$. Максимального значення індекс I досягає, коли експертів розділяють порівну на дві групи, виставляючи крайні оцінки зі шкали оцінювання, тобто якщо оцінювання проводять на шкалі $[0, s]$, а в оцінюванні беруть участь n експертів, то $I = 0$, коли множина оцінок має вигляд $\left\{\frac{1 \dots 1}{2}, \frac{s \dots s}{2}\right\}$ за парного n та $\left\{\frac{1 \dots 1}{2}, \frac{s \dots s}{2}\right\}$ або $\left\{\frac{1 \dots 1}{2}, \frac{s \dots s}{2}\right\}$ за непарного n .

4.2. Метод визначення попарної узгодженості експертних оцінок

Концепцію запропонованого у роботі [4] методу можна використовувати для підрахунку індексів узгодженості між оцінками емоційності коментарів, пов'язаних з новинними подіями з різних ІР. Спершу визначимо термінологічні аналогії класичної задачі підтримки прийняття рішень і задачі визначення напруженості в суспільстві. По-перше, роботу експертів, які у класичній задачі оцінюють альтернативні рішення за певною шкалою, у задачі визначення напруженості виконують нейронні мережі, що визначають емоційність коментарів. Утім, формат експертних оцінок суттєво відрізняється: якщо у класичній задачі оцінки — це числові значення з певної заздалегідь визначеної множини, що показують у попарному порівнянні, наскільки одна альтернатива краща за іншу, то у задачі визначення напруженості кожна оцінка — це дискретний розподіл вагових коефіцієнтів для оцінок напруженості на інтервалі $[0, 2]$. По-друге, альтернативам у класичній задачі відповідає вплив деякої новинної події на підвищення рівня напруженості у різних ІР.

Щоб використати запропонований авторами роботи [4] метод для задачі визначення рівня

напруженості, треба відкоригувати формулу (6). При цьому основний зміст вимог до вихідної формули має бути збережений.

На відміну від точкових оцінок, які порівнюють у класичному методі, в адаптованому методі порівнюють пари оцінок з урахуванням ваг. Початково ваги ненормовані, тому спершу потрібно нормалізувати ваги у межах кожного ІР для певної новини за формулою

$$p_i = \frac{w_i(x)}{\sum_z w_i(z)}, \quad (7)$$

де w_i – ваги, що відповідають окремим оцінкам емоційності коментарів; $x \in X$, X – множина дискретних оцінок на шкалі оцінювання коментарів; $\sum_z w_i(z)$ – сума всіх ваг для досліджуваної новини (кількість коментарів, пов'язаних із новинною подією).

Щоб врахувати попарні відмінності, треба масштабувати кожну пару оцінок добутком їхніх нормованих ваг. Вважатимемо також, що $f(x) = |x|$, тоді

$$\delta_{ij} = \frac{1}{R} \sum_x \sum_y p_i(x) p_j(y) |x - y|, \quad (8)$$

де δ_{ij} – парна «несхожість» між оцінками новини на різних ІР; $p_i(x)$, $p_j(y)$ – ваги порівнюваних оцінок новини із двох ІР; x , y – оцінки новин; $R = \max(x) - \min(x)$ – різниця між верхньою та нижньою границями шкали оцінювання. Подвійна сума у формулі (8) досягатиме свого максимуму, коли на одному ІР всі коментарі будуть мати максимальну негативну емоційність, а на іншому – максимальну позитивну. Тоді єдиним ненульовим членом подвійної суми буде елемент з $p_i(x) = 1$ та $p_j(y) = 1$, а $|x - y| = R$, що в результаті дасть найбільше можливе значення $\delta_{ij} = 1$. Через нормування ваг можна відійти від суми модулів різниць абсолютних значень оцінок M , оскільки верхню межу вже закладено у самій формулі.

Усі вимоги, що ставились до індексу узгодженості, заданого формулою (6), виконуються для модифікованих формул (7)–(8).

4.3. Визначення умови фільтрації ресурсів з низьким рівнем узгодженості

Рівень узгодженості вимірюють з метою перевірки можливості проведення агрегації оцінок впливів новинних подій з різних ІР. Якщо рівень узгодженості оцінок для новин з деяких ресурсів суттєво нижчий за інші, ці оцінки не будуть використані під час агрегації, натомість їх вра-

ховуватимуть безпосередньо у загальному внеску новинної події у підвищення рівня СН. Щоб відфільтрувати оцінки, які мають низький рівень узгодженості порівняно з іншими, задають порогові значення.

Перш ніж починати досліджувати рівень узгодженості окремих ІР варто перевірити загальний середній рівень узгодженості між усіма ІР, тобто визначити загальний ступінь поляризації думок. Якщо загальний ступінь поляризації буде високим, то агрегація не буде давати показових результатів.

На основі оцінок δ_{ij} для кожної пари ІР будують матрицю $D_{ij} \in [0, 1]^{K \times K}$; $D_{ij} = 0$. Підрахунок загальної поляризації M проводять за формулою

$$M = \frac{2}{K(K-1)} \sum_{i < j} D_{ij}, \quad (9)$$

де K – кількість досліджуваних ІР, $M \in [0, 1]$.

Якщо припустити, що всі ІР розділились на дві групи із частками π та $1 - \pi$, де π – це мінімальна частка меншої групи, то порогове значення M_t можна розрахувати за формулою

$$M_t = 2\pi(1 - \pi)\delta_{\min}, \quad (10)$$

де $\delta_{\min} \in [0, 1]$ – коефіцієнт мінімальної міжгрупової несхожості, $M_t \in [0, 1]$. Цей коефіцієнт показує мінімальну допустиму несхожість між парою ІР, яку вважатимуть суттєвою для того, щоб визначати думки читачів цих ресурсів як поляризовані. Значення π теж обирають емпірично, воно вказує на мінімальну частку групи ІР, оцінки яких суттєво різняться від інших. Якщо для якоїсь новинної події загальна поляризація оцінок різних ІР більша за порогове значення ($M > M_t$), агрегацію для цієї події не проводять, якщо ж поляризація менша, то проводять аналіз узгодженості усіх ІР.

Щоб визначити середню парну несхожість кожного ресурсу з рештою, розраховують середні значення за рядками матриці D_{ij} за формулою

$$\mu_i = \frac{1}{K-1} \sum_{j \neq i} D_{ij}. \quad (11)$$

Ідея методу визначення неузгоджених ресурсів полягає у тому, щоб дослідити, як вплине на загальне значення індексу узгодженості. Якщо значення індексу узгодженості після цього суттєво зросте, це означатиме, що цей ІР мав досить низьку узгодженість з іншими. Оцінки, отримані на основі аналізу цих ІР, не будуть агрегуватись разом з рештою оцінок.

Індекс узгодженості I загалом може бути визначений як $1 - M$. Тоді, якщо позначити через $I_{(-i)}$ значення індексу узгодженості після вилучення ресурсу i , а через ΔI_i зміну індексу узгодженості після цього вилучення, то справедливим буде вираз

$$\Delta I_i = I_{(-i)} - I = (1 - M_{(-i)}) - (1 - M) = M - M_{(-i)}. \quad (12)$$

Якщо вважати, що $\sum_{i < j} D_{ij} = T$, а $\sum_{i \neq j} D_{ij} = s_i$, то

$$M = \frac{2T}{K(K-1)}, \quad (13)$$

$$M_{(-i)} = \frac{2(T - s_i)}{(K-1)(K-2)}. \quad (14)$$

Отже, з формул (12)–(14) можна отримати формулу

$$\begin{aligned} \Delta I_i &= \frac{2T}{K(K-1)} - \frac{2(T - s_i)}{(K-1)(K-2)} = \\ &= \frac{2}{K-2} \left(\frac{s_i}{K-1} - \frac{2T}{K(K-1)} \right) = \\ &= \frac{2}{K-2} (\mu_i - M), \end{aligned} \quad (15)$$

де $K > 2$.

Якщо $\Delta I_i > 0$, це означає, що вилучення ресурсу підвищує індекс узгодженості, отже, його узгодженість з рештою ресурсів нижча за середнє значення. Утім, для вилучення ресурсу з підрахунку агрегованої оцінки потрібно підібрати додаткове порогове значення, яке дозволить відфільтрувати ресурси з низьким рівнем узгодженості. Для цього введемо поріг τ як частку від оцінки загальної поляризації, яку розраховують за формулою

$$\tau = \frac{2}{K-2} \eta M, \quad (16)$$

де $\eta \in [0, 1]$ — коефіцієнт мінімально значущої частки, що показує мінімальну частку поляризації, яку вважають практично значущою.

Якщо для деякого ресурсу i приріст узгодженості після його вилучення досягне або перевищить значення визначеного порогу ($\Delta I_i > \tau$), то такий ресурс слід вилучати з підрахунку агрегованої оцінки впливу досліджуваної новинної події.

5. Метод коригування індексу соціальної напруженості на основі методів визначення узгодженості експертних оцінок

Підрахунок загального індексу соціальної напруженості без урахування узгодженості між ресурсами може призвести до втрати інформативності значень індексу. Для агрегації оцінок впливу різних новинних подій беруть лише узгоджені між собою оцінки, натомість значення неузгоджених враховують окремо у загальній оцінці впливу новинної події.

Для врахування оцінок неузгоджених ресурсів розглянемо два випадки: 1) коли загальна поляризація в контексті окремої новинної події вища за порогове значення ($M > M_t$), 2) коли рівень узгодженості окремих ІР з іншими досить низький (ресурси i , для яких $\Delta I_i > \tau$). Зауважимо, що ці два випадки вважають такими, що не можуть виконуватись одночасно, тобто якщо рівень поляризації для деякої новинної події перевищує порогове значення, то рівень узгодженості окремих ІР не підраховують, і навпаки, якщо порогове значення для поляризації виконується, проводять дослідження кожного окремого ІР щодо його узгодженості з рештою ресурсів.

Розглянемо перший випадок. Якщо загальна поляризація M_e перевищує поріг M_t , слід визначити, наскільки великим є це перевищення. Для цього розраховують нормований індекс перевищення порога P_e для події e за формулою

$$P_e = \frac{M_e - M_t}{1 - M_t}, \quad (17)$$

де $P_e \in [0, 1]$. Нормований індекс P_e показує частку досягнутого перевищення порога від максимально можливого перевищення.

Якщо загальна поляризація M_e перевищує поріг M_t , агрегацію впливів новинних подій на підвищення СН не підраховують за формулою (5). У цьому випадку застосовують формулу (18) для $\tilde{v}'_e$, в якій використовують індекс перевищення порога P_e , зважений коефіцієнтом глобального підсилення α :

$$\tilde{v}'_e = \alpha P_e, \quad (18)$$

де α — коефіцієнт глобального підсилення, $\alpha \in [0, 1]$.

Розглянемо другий випадок. Якщо загальна поляризація M_e задовольняє поріг M_t ($M_e \leq M_t$), але деякі ресурси мають низький рівень узгодженості з рештою, спочатку для цих ресурсів підраховують локальний нормований надлишок неузгодженості $s_{e,i}$ за формулою

$$s_{e,i} = \frac{\Delta I_{e,i} - \tau_e}{\frac{2}{K-2}(1 - M_e)}, \quad (19)$$

де $s_{e,i} \in [0, 1]$, $K > 2$. Він показує, наскільки сильно ресурс i для події e перевищує поріг неузгодженості. У чисельнику підраховують відхилення значення $\Delta I_{e,i}$ від τ_e , а у знаменнику — максимально можливе значення $\Delta I_{e,i}$.

Далі знаходять середнє значення C_e локальних нормованих надлишків неузгодженості $s_{e,i}$ для тих ресурсів, які мають низький рівень узгодженості з рештою, за формулою

$$C_e = \frac{1}{K-1} \sum_i s_{e,i}, \quad (20)$$

де $C_e \in [0, 1]$. Отже, що більше буде неузгоджених ІР, то більшим буде значення C_e .

Значення C_e використовують для коригування ступенів досягнення цілей другого рівня, яке проводять за формулою

$$\tilde{v}'_e = \tilde{v}_e + (1 - \tilde{v}_e)\beta C_e, \quad (21)$$

де β — коефіцієнт локального підсилення, $\beta \in [0, 1]$.

Після коригування внесків у підвищення СН тих подій, для яких оцінки новинних ресурсів не мають повної узгодженості, перераховують індекс СН, який під впливом змінених внесків новинних подій теж змінює значення і набуває більшої інформативності.

6. Приклад підрахунку скоригованого індексу соціальної напруженості на основі методу узгодженості експертних оцінок

6.1. Визначення коефіцієнтів для моделі

Для практичного застосування розробленої математичної моделі визначення СН треба визначити константні параметри (коефіцієнти), що їх використовують для підрахунку порогових значень та калібрування чутливості індексів P_e та C_e . До цих параметрів входять:

- $\delta_{\min}$ — коефіцієнт мінімальної міжгрупової несхожості;
- π — параметр мінімальної частки меншої групи у моделі з двома групами;
- η — коефіцієнт мінімально значущої частки, що показує мінімальну частку поляризації, яку вважають практично значущою;
- α — коефіцієнт глобального підсилення, коли $M_e > M_i$;

• β — коефіцієнт локального підсилення, коли $M_e \leq M_i$, і для деяких ІР $\Delta I_i > \tau$.

Коефіцієнт $\delta_{\min}$ і параметр π використовують для підрахунку порогового значення M_i індексу загальної поляризації M_e за формулою (10). Поріг M_i повинен визначати мінімальний рівень поляризації, який має допускати можливість певного відхилення в оцінках деяких ІР, і одночасно бути достатньо жорстким, щоб відслідковувати поділ оцінок ресурсів на дві групи.

У межах практичної апробації моделі пропонуємо взяти $\delta_{\min} = 0,7$. На шкалі оцінювання емоційності $[0, 2]$ обране значення $\delta_{\min}$ буде відповідати оцінці 1,4, тобто відстань між центрами груп на шкалі $[0, 2]$ має бути меншою за 1,4, щоб не перевищувати порогового значення. Це обмеження є досить жорстким, але коли мова йде саме про глобальну поляризацію, воно відповідає очікуванням щодо обмежень.

Значення параметра мінімальної частки меншої групи π пропонуємо взяти рівним 0,25. Він показує, що оцінки будемо вважати поляризованими у тому випадку, якщо щонайменше чверть ІР потрапить в одну групу, а решта — в іншу. Якщо досліджуватимемо, наприклад, 10 ІР, то новинну подію будемо вважати такою, що має поляризовані думки, якщо принаймні для 3 із 10 ІР центр оцінок буде відрізнятися не менше, ніж на 1,4, порівняно з центром решти ресурсів.

Тоді за $\delta_{\min} = 0,7$ й $\pi = 0,25$ відповідно до формули (10) значення порогу M_i буде дорівнювати 0,2625, що відповідає очікуваній жорсткості.

Коефіцієнт мінімально значущої частки η використовують для підрахунку порогу зростання індексу узгодженості з вилученням деякого ресурсу у процесі підрахунку цього індексу. З формул (15)–(16), використавши елементарні перетворення, можна вивести умову

$$\mu_i \geq M(1 + \eta). \quad (22)$$

Отже, $(1 + \eta)$ визначає коефіцієнт, що показує, у скільки разів середня несхожість ресурсу з іншими ресурсами має бути більшою за середню несхожість за всіма ресурсами.

Пропонуємо взяти значення η рівним 1, тобто ресурс будемо вважати таким, що суттєво відрізняється за оцінками емоційності від інших, якщо його середня несхожість з іншими ресурсами буде принаймні вдвічі більшою за середню несхожість за всіма ресурсами.

Коефіцієнт глобального підсилення α визначає ступінь впливу індексу перевищення порога P_e на агрегований внесок події у підвищен-

ня СН $\tilde{v}_e$, який підраховують за формулою (18) у випадку, коли $M_e > M_r$. Оскільки тут розглядаємо великий ступінь поляризації, пропонуємо взяти $\alpha = 1$. Це дасть можливість за певних обставин агрегованому внеску події $\tilde{v}_e$ досягати значення 1, що має бути допустимим з огляду на можливість максимальної поляризації.

Коефіцієнт локального підсилення β дозволяє підвищити $\tilde{v}_e$ за рахунок оцінок подій, які не використовували під час агрегації через низьку узгодженість із рештою оцінок. Пропонуємо взяти $\beta = 1$. Оскільки C_e майже завжди буде виходити значно ближчим до 0, ніж до 1, коефіцієнт β доцільно взяти максимально великим.

6.2. Приклад підрахунку індексу соціальної напруженості без урахування узгодженості

Для проведення практичного експерименту було змодельовано ситуацію, в якій досліджували рівень соціальної напруженості за даними з 10 ІР. Ці дані являють собою розподіли оцінок коментарів до новинних публікацій, які стосуються 10 новинних подій, причому кожний з розглянутих ІР висвітлив кожну з цих подій у кількох публікаціях. Побудована база знань СППР має мережеву структуру, тому одній новинній публікації можуть відповідати 1–3 новинні події, а кожній новинній події можуть відповідати 5–20 публікацій у межах кожного ІР. Внесок новинної публікації у підвищення СН характеризується оцінками емоційності її коментарів, розподілених на інтервалі $[0,2]$ із кроком 0,5, починаючи від нуля. Загалом було змодельовано унікальні розподіли оцінок емоційності для 940 новинних публікацій на 10 ІР. На основі цих розподілів за алгоритмом розрахунку індексів соціальної напруженості були підраховані внески кожної події у підвищення СН у межах кожного ІР та загальні внески за всіма ІР, а також підраховано ступінь досягнення головної цілі, який інтерпретуємо як загальний рівень СН. Оскільки всі підрахунки проводили в один момент часу t , значення активності користувачів вважали однаковим для всіх ІР й подій. У табл. 1 наведено перелік усіх новинних подій та їх агреговані внески $\tilde{v}_e$ у підвищення СН.

Якщо підраховувати суму всіх значень $\tilde{v}_e$ за формулою (4), то вийде, що для випадку, коли рівень узгодженості між ресурсами не враховували, рівень СН досягає значення $\Theta = 0,857\,143$.

Таблиця 1. Новинні події та їхні агреговані внески $\tilde{v}_e$ у підвищення соціальної напруженості

Порядковий номер події	Агреговані внески подій $\tilde{v}_e$
1	0,084 505
2	0,071 029
3	0,042 097
4	0,036 288
5	0,143 665
6	0,163 230
7	0,157 683
8	0,158 646

6.3. Приклад підрахунку індексу соціальної напруженості з урахуванням узгодженості

Щоб забезпечити об'єктивність експерименту, у цьому прикладі розглянемо той самий набір вхідних даних, що і в попередньому прикладі. Утім, агрегований внесок події у підвищення СН підраховуємо за формулою (5) лише тоді, коли загальний рівень поляризації між ІР в межах новинної події M_e не перевищує порогового значення M_r , а приріст узгодженості для кожного ІР не перевищує порога τ . Якщо не виконується умова $M_e \leq M_r$, агрегований внесок $\tilde{v}'_e$ події визначають за формулою (18), якщо ж ця умова виконується, але для деяких ІР у межах новинної події відбувається перевищення приросту узгодженості після вилучення цього ІР ($\Delta I_i > \tau$), тоді агрегований внесок $\tilde{v}'_e$ події знаходять за формулою (21). Розглянемо табл. 2 з результатами експерименту з урахуванням узгодженості, а також результатами попереднього експерименту для порівняння.

У табл. 2 відображено номери подій (№ з/п), загальні рівні поляризації по кожній події (M_e), середні значення локальних нормованих надлишків неузгодженості (C_e), нормовані індекси перевищення порога у випадку високої глобальної поляризації (P_e), агреговані внески подій без урахування узгодженості ($\tilde{v}_e$) та агреговані внески подій з урахуванням узгодженості ($\tilde{v}'_e$).

З табл. 2 видно, що значення загального рівня поляризації для подій 1-4 вищі за інші, отже, потенційно внески цих подій у підвищення СН мають бути вищими. Утім, значення агрегованих внесків подій без урахування узгодженості цього не відображають. Навпаки, значення $\tilde{v}_e$ для подій 1-4 нижчі, ніж для подій 5-8. Такий результат

Таблиця 2. Таблиця результатів експерименту з підрахунку індексу соціальної напруженості з урахуванням узгодженості

№ з/п	M_e	C_e	P_e	$\tilde{v}_e$	$\tilde{v}'_e$
1	0,55 314 824	0	0,3 940 993 084	0,08 450 464 627	0,3 940 993 084
2	0,5 315 935 961	0	0,3 648 726 727	0,07 102 860 226	0,3 648 726 727
3	0,4 675 525 378	0	0,2 780 373 394	0,04 209 736 551	0,2 780 373 394
4	0,3 615 710 597	0	0,1 343 336 403	0,03 628 827 511	0,1 343 336 403
5	0,1 654 205 798	0	0	0,1 436 654 445	0,1 436 654 445
6	0,1 723 429 566	0	0	0,1 632 295 144	0,1 632 295 144
7	0,1 235 095 482	0	0	0,1 576 834 506	0,1 576 834 506
8	0,1 254 018 918	0	0	0,1 586 455 585	0,1 586 455 585

є наслідком втрати інформативності оцінок під час підрахунків за формулами (3)–(5). Щоб підвищити внески подій з вищим рівнем поляризації, застосовують індекси P_e і C_e . Вище було встановлено, що оптимальним порогом для загальної поляризації буде значення $M_t = 0,2625$. Оскільки для подій 1-4 виконується умова $M_e > M_t$, для цих подій підраховують індекси перевищення порога P_e , які, за $\alpha = 1$, дорівнюють оновленим агрегованим внескам подій $\tilde{v}'_e$. Індекси C_e можна розраховувати лише коли $M_e \leq M_t$, тобто лише для подій 5-8, але оскільки порогового значення несхожості жодного ІР у межах жодної з цих чотирьох подій не було досягнуто, індекс C_e залишився рівним нулю для всіх подій.

Дані табл. 2 свідчать, що агреговані внески подій з урахуванням узгодженості для подій з високою поляризацією істотно перевищують відповідні внески без урахування узгодженості. Окрім того, порівняно з експериментом без урахування узгодженості оцінка рівня СН виросла до $\Theta = 1,794567$, що пояснюється врахуванням високої поляризації оцінок ресурсів для деяких новин.

Загалом результати показують дієвість обраного методу і правильність підбору коефіцієнтів для експерименту.

Висновки

1. Запропоновано метод застосування системи підтримки прийняття рішень для визначення індексу СН. Розроблено алгоритм розрахунку індексу СН з використанням математичного апарату систем підтримки прийняття рішень.

2. Проаналізовано відомі методи визначення узгодженості експертних оцінок і запропоновано модифікований метод.

3. Розроблено метод коригування індексу соціальної напруженості на основі методів визначення узгодженості експертних оцінок.

4. Запропоновано методику визначення коефіцієнтів для підрахунку порогових значень поляризації та зміни індексу узгодженості, а також коефіцієнтів локального та глобального підсилення для підрахунку нових агрегованих внесків подій у підвищення СН.

5. Проведено експериментальне дослідження ефективності методу визначення індексу СН із застосуванням методу визначення узгодженості експертних оцінок. Із застосуванням цього методу зафіксовано підвищення індексу соціальної напруженості з 0,857 143 до 1,794 567 за високого рівня поляризації ресурсів у межах деяких новинних подій.

Наступні дослідження будуть присвячені покращенню розробленої математичної моделі СН, її тестуванню на реальних даних, пошуку способів більш гнучкого налаштування параметрів цієї моделі.

References

- [1] O. Rudachenko and T. Klebanova, "Suchasni pidkhody do analizu sotsialnoi napruzhenosti", *Naukovyi visnyk Uzhhorodskoho natsionalnogo universytetu. Seriya: Mizhnarodni ekonomichni vidnosyny ta svitove hospodarstvo*, 2020, no. 30, pp. 140–144.
- [2] O. Andriichuk and M. Shchoholiev, "Zastosuvannia metodiv pidtrymky pryiniattia rishen ta obchysluvalnoi linhvistyky dlia vyznachennia rivnia sotsialnoi napruzhenosti", *Naukovi notatky*, 2023, pp. 217–224.
- [3] M. Shchoholiev and O. Andriichuk, "Matematychni metody obrobky pryrodnoi movy u systemi operatyvnoho vyznachennia rivnia napruzhenosti v suspilstvi", *Kibernetyka ta kompiuterni tekhnolohii*, 2023, no. 2, pp. 55–68.
- [4] V.V. Tsyhanok and P.D. Roik, "Metod vyznachennia ta pidvyshchennia uzghodzhennosti ekspertnykh otsinok za pidtrymannia pryiniattia hrupovykh rishen", *System Research and Information Technologies*, 2018, no. 3, pp. 110–121.
- [5] F.L. Schmidt and J.E. Hunter, "Interrater reliability coefficients cannot be computed when only one stimulus is rated", *Journal of Applied Psychology*, 1989, vol. 74, pp. 368–370.
- [6] M.J. Burke *et al.*, "On average deviation indices for estimating interrater agreement", *Organizational Research Methods*, 1999, vol. 2, pp. 49–68.
- [7] L.R. James *et al.*, "Estimating within-group interrater reliability with and without response bias", *Journal of Applied Psychology*, 1984, vol. 69, pp. 85–98.
- [8] M.K. Lindell *et al.*, "A revised index of agreement for multi-item ratings of a single target", *Applied Psychological Measurement*, 1999, vol. 23, pp. 127–135.
- [9] R.D. Brown and N.M.A. Hauenstein, "Interrater agreement reconsidered: An alternative to the rwg indices", *Organizational Research Methods*, 2005, vol. 8, pp. 165–184.
- [10] M.Z. Zgurovsky *et al.*, "Group incomplete paired comparisons with account of expert competence", *Mathematical and Computer Modelling*, 2004, vol. 39, no. 4–5, pp. 349–361.
- [11] G. A new measure of rank correlation, *Biometrika*, 1938, vol. 30, no. 1–2, pp. 81–93. Retrieved from: doi: 10.1093/biomet/30.1-2.81.
- [12] A. Olenko and V. Tsyganok, "Double entropy inter-rater agreement indices", *Applied Psychological Measurement*, 2016, vol. 40, no. 1, pp. 37–55.

M.O. Shchoholiev, V.V. Tsyganok

A METHOD FOR ADJUSTING THE SOCIAL TENSION INDEX BASED ON METHODS FOR ASSESSING THE CONSISTENCY OF EXPERT ESTIMATES

Problem. Social tension is the subject of numerous contemporary sociological studies due to its significant impact on the effectiveness of decision-making both at the governmental level and in business contexts. Social tension can be assessed by analyzing public reactions to news events in social networks. However, obtaining generalized estimates of the impact of news events on increasing levels of social tension is complicated by the fact that audiences of different information channels may hold significantly different opinions about the same event. The applicability of methods for assessing the consistency of expert estimates to problems that do not involve traditional expert input remains insufficiently studied.

Objective. To develop and test a method for adjusting the social tension index using methods for assessing the consistency of expert estimates.

Methodology of implementation. The implementation of the method for adjusting the social tension index involves the interaction between the knowledge base of a decision support system and a method for assessing the consistency of expert estimates, based on minimizing the total distance between all pairwise estimates.

Results. A method for adjusting the social tension index based on methods for assessing the consistency of expert estimates is proposed. A procedure for determining a threshold consistency level for evaluating the impact of news events on social tension is introduced. Applying the proposed method for adjusting the social tension index makes it possible to improve the validity of social tension assessments.

Conclusions. The application of the proposed method for adjusting the social tension index based on methods for assessing the consistency of expert estimates enhances the validity of social tension assessments in cases where the consistency index among estimates of news event impacts falls below the defined threshold.

Keywords: Social tension; social networks; news event; decision support system; consistency of estimates; an aggregation-adequate level of consistency.

Рекомендована Радою
факультету прикладної математики
КПІ ім. Ігоря Сікорського

Надійшла до редакції
22 червня 2025 року

Прийнята до публікації
08 вересня 2025 року