



ІНФОРМАЦІЙНА ТА КІБЕРНЕТИЧНА БЕЗПЕКА

УДК 004.621.68

DOI 10.17721/ISTS.2023.1.5-12

Микола Браїловський, orcid.org/0000-0002-3148-1148,

bk1972@ukr.net

Київський національний університет імені Тараса Шевченка, Київ, Україна,

Володимир Хорошко, orcid.org/0000-0001-6213-7086,

professor_va@ukr.net

Національний авіаційний університет, Київ, Україна

МЕТОДИ ДЕСТРУКТИВНОГО ВПЛИВУ ТА ЗАХИСТУ КОНТЕНТУ В СОЦІАЛЬНИХ МЕРЕЖАХ

Розглянуто актуальну тему захисту суспільства від деструктивних впливів у соціальних мережах. Доведено, що для запобігання подібним впливам, необхідно оцінювати не тільки матеріали, що містяться в текстах, аудіо- та відеоконтентах, а й тональність інформації, що поширюється в них. Тональність є однією з характеристик думки або емоції і нині використовується як основний параметр для класифікації текстів за тематичними категоріями. Це пояснюється зручністю й ефективністю класифікації за цим принципом та значно спрощує інші завдання з класифікації текстових постів. Сформовано категорії тональної оцінки. Доведено можливість оцінювання настрою за одновимірним або багатовимірним параметром. Показано можливість використання для класифікації тональності тексту двох підходів. Перший – це автоматичний аналіз цифрових текстів комп'ютерами з елементами машинного навчання або навчання з учителем. Другий – за допомогою словників (тезаурусів), підключених до програми. Наведено недоліки, які виникають у автоматичних комплексах для визначення тональності. Запропоновано новий алгоритм роботи автоматичних комплексів. Зазначено, якщо якість системи аналізу тональності тексту оцінюється тим, наскільки добре вона узгоджується з думкою людини щодо емоційного оцінювання досліджуваного тексту, то для цього можуть використовуватися такі показники, як точність і повнота. Наведено розрахунки цих параметрів. Під час розгляду публікації, представленої у формі зображення, зазначено необхідність враховувати параметри яскравості, контрасту, а також колірних поєднань, оскільки ці показники потенційно можуть використовуватися розповсюджувачами деструктивного контенту для привернення уваги мережі користувачів. Заявлено, що фотоматеріали можуть містити різні праворадикальні знаки та символи, які можуть стосуватися пропаганди певної забороненої організації чи ідеології. Представлено частоти, завдяки яким можливо дестабілізаційно впливати на організм людини. Проведено аналіз можливих методів впливу на свідомість користувача соціальними мережами.

Ключові слова: деструктивні впливи аудіо- та відеоконтентів; соціальні мережі; аналіз тональності тексту; методи аналізу.

1. ВСТУП

Соціальні мережі нині відіграють важливу роль у житті кожної людини. Для них немає кордонів, тому вони об'єднують аудиторію не лише жителів однієї держави, а й усієї планети. За їх допомогою обмінюються різноманітними типами контенту: текстовим, аудіо, фото та відео, що у свою чергу серйозно впливає на світогляд кожної окремої

людини. У зв'язку із цим виникає необхідність захисту соціально-інформаційного простору шляхом виявлення деструктивного контенту, який є шкідливим для користувачів та суспільства в цілому [1]. Для запобігання подібним впливам, необхідно оцінювати не тільки матеріали, що містяться в текстах, аудіо- та відеоконтентах, а й тональність інформації, що поширюється в них [2, 3].

© Браїловський М., Хорошко В., 2023



2. ВИКЛАДЕННЯ ОСНОВНОГО МАТЕРІАЛУ

Аналіз тональності тексту – клас методів контент-аналізу в комп'ютерній лінгвістиці, призначених для автоматичного виявлення емоційно забарвленої лексики в текстах та емоційної оцінки авторів щодо об'єктів, про які йдеться у тексті.

Тональність – емоційне ставлення автора висловлювання до якогось об'єкта, виражене в тексті. Емоційний компонент, виражений на рівні лексеми чи комунікативного фрагмента, називається лексичною тональністю. Тональність усього тексту в цілому можна визначити як функцію лексичної тональності складових його одиниць і правил їхнього поєднання [4].

Тональність є однією з характеристик думки або емоції і нині використовується як основний параметр для класифікації текстів за тематичними категоріями. Це пояснюється зручністю й ефективністю класифікації за цим принципом та значно спрощує інші завдання щодо класифікації текстових постів. Зручність полягає у спрощенні класифікації шляхом зведення її до оцінювання лише одного параметра. Результат визначення класу тексту за настроєм можна уточнити за допомогою інших зазвичай відомих параметрів (наприклад, автора публікації, регіону, в якому вона була написана, спільноти користувачів тощо). Тематичні категорії для класифікації слід визначити заздалегідь. Автор і тема публікації можуть бути відомі, і в цьому випадку головною метою аналізу є визначення їхнього відношення до обговорюваної теми. Якщо автор та/або тема невідомі, то першим завданням є визначення теми обговорення, наступним завданням є оцінювання тональності висловлювань автора, а потім, за потреби, визначення автора публікації [5, 6].

Незважаючи на те, що завдання класифікації зводиться, здавалося б, до одновимірності (визначення одного параметра тексту – його тональності), воно залишається досить складним, оскільки сам параметр може бути як одновимірним, так і векторним, і необхідно вибрати шкалу і правила відповідності її значень. Багатовимірний параметр оцінюється за кількома шкалами. Отже, тональність може відображати як одновимірний емоційний простір, так і багатовимірний. Якщо для відображення емоцій обрано одновимірний простір, то параметр "тональність" може набувати одне з можливих значень n -вимірної шкали. Якщо вибрано багатовимірний простір, то необхідно вибрати основу для цього простору, тобто базові емоції –

радість, гнів, страх тощо [5], які перетворюються на шкали оцінок.

Аналіз тональності проводять для того, щоб, по-перше, визначити ставлення автора до предмета обговорення, по-друге, виявити авторську позицію, по-третє, виявити додаткові ознаки думки (наприклад, первинна вона чи запозичена, чи є подібні погляди). Так, у [7] думки поділено на два типи: пряма думка і порівняння.

Безпосередню думку можна формально представити як кортеж [7]:

$$O = (e, f, op, h, t), \quad (1)$$

де e (*сутність*) – об'єкт настрою, тобто предмет обговорення; f (*feature*) – властивості об'єкта тональності (атрибути, частини об'єкта); op (*орієнтація* або *полярність*) – тональна оцінка (емоційна позиція автора щодо згаданої теми); h (*тримач*) – суб'єкт тональності (автор думки); t (*час*) – момент часу, коли була залишена думка.

Порівняння здійснюють відносно будь-яких предметів, явищ, подій або їхніх властивостей. Порівнюють дві або більше думок, кожна з яких може бути представлена як первинна (1). Результатом порівняння є думка про перевагу одного об'єкта над іншим (перевага властивості одного об'єкта над подібною властивістю іншого об'єкта).

Ознакою переваги в тексті є ступені порівняння прикметників, частин мови та морфем (дієслова, прислівники, суфікси тощо). Очевидно, що для різних мов необхідно формувати відповідні набори ознак.

Формальне представлення порівняння залежить від його мети. Наприклад, якщо порівнюють два об'єкти, то формальне представлення може бути таким [7]:

$$C = (O_1, O_2, F, Op, h, t), \quad (2)$$

де O_1, O_2 – об'єкти порівняння; F – набір загальних характеристик, які порівнюються; Op – сукупність уподобань автора думки; h – суб'єкт тональності (автор, тобто той, кому належить ця думка); t – момент часу, в який висловлено порівняльну думку.

Тональну оцінку можна сформувати за такими категоріями [3, 6–8]:

- 1) позитивна, негативна;
- 2) позитивна, негативна, нейтральна (відсутність емоційного забарвлення);
- 3) більше градацій, наприклад, позитивна, середньопозитивна, злегка позитивна, нейтральна, злегка негативна, середньонегативна, негативна.

Оцінювання настрою може здійснюватися за одновимірним або багатовимірним параметром [5, 6, 8].



1. *Класифікація тональності за бінарною шкалою.* Бінарна шкала передбачає дві оцінки: позитивну і негативну. Це найпростіший спосіб класифікації з погляду формалізації завдання, якщо висловлена однозначна думка. Проте автор може висловити свою точку зору, наприклад, аналізуючи об'єкт із різних сторін. Думка може містити ознаки як позитивного, так і негативного емоційного забарвлення. У цьому випадку однозначно оцінити тональність дуже складно і не завжди можливо.

2. *Класифікація тональності за багатомірною шкалою.* Між полярними оцінками, позитивними та негативними, вводять додаткові градації. Чим більше градацій, тим складніше завдання класифікації. Потрібна поступова класифікація.

Етап 1. Класифікація нейтральної тональності.

Етап 2. Класифікація негативної/позитивної тональності.

Етап 3. Класифікація ступеня позитивності/негативності.

3. *Системи градації (шкалування)* формують шляхом надання різним словам відповідного числа з вибраного інтервалу, наприклад, від мінус 1 до 1. При цьому мінус 1 відповідає найбільш негативній тональності, 0 – нейтральній, а 1 – найбільш позитивній. Тобто кожному слову присвоюється ступінь важливості. Це, по суті, і є формування вектора ваг. Текст, який аналізується, попередньо досліджується за допомогою засобів оброблення мови (інструментів і алгоритмів) з метою виокремлення слів або фрагментів для подальшого оцінювання тональності.

4. *Суб'єктивність/об'єктивність.* Визначення суб'єктивності/об'єктивності тексту – окрема проблема. Складність посилюється тим, що слова або фрази можуть використовуватися в певному контексті. Крім того, слід розрізняти, коли автор цитує чужу думку, з якою він може повністю або частково погоджуватись чи не погоджуватись. Об'єктивним вважається "неупереджений" текст, який не містить судження, а тільки викладає поточний стан речей. Наприклад, стаття новин. Ідентифікацію об'єктивності/суб'єктивності використовують як додатковий етап класифікації змісту: попередньо вилучивши об'єктивні думки (речення), можна отримати точнішу оцінку настрою.

Детальніший аналіз можна провести на основі моделі функції/аспекту. Спочатку встановить об'єкти (властивості), за якими буде оцінюватися тональність, а потім визначте її якимось чином (наприклад, як функцію). У процесі розв'язування задачі класифікації визначається об'єкт, виокремлюється відповідна йому функція та оцінюється тональність.

Для класифікації тональності тексту використовують два підходи. Перший – це автоматичний аналіз цифрових текстів комп'ютерами з елементами машинного навчання або навчання з учителем [9, 10]. Другий – за допомогою словників (тезаурисів), підключених до програми [11–15]. Відмінності між цими підходами очевидні: перший забезпечує більш точний результат і більший обсяг даних для оброблення, другий – легший у реалізації. Якщо за допомогою другого підходу вдається не повністю завантажувати базу даних, а оптимізувати використання словників, то швидкість оброблення й обсяг оброблених даних можна збільшити.

Усі згадані вище тезауруси були створені для англійської мови. Для української мови подібних розробок практично немає. Тому розглянемо, як приклад, подібні твори для російської мови.

Зараз у Санкт-Петербурзькому державному університеті кафедра математичної лінгвістики розробляє базу даних російською мовою RussNet, подібну до WordNet. Тезаурус RussNet містить 1300 іменників, 1900 дієслів, 1100 прикметників, 200 прислівників. Внутрішньомовна структура мови недостатньо розвинена, тому тезаурус не пов'язаний з Inter-Lingual-Index [16].

Програма Абріаль [17] була розроблена 2004 р. А. І. Пацкіним у Російському науково-дослідному інституті штучного інтелекту з метою аналізу текстів російською мовою. Програма застосовна для широкого кола завдань, включаючи аналіз настроїв, але має істотний недолік: база даних повністю завантажується у віртуальну пам'ять. Тезаурус Роже був використаний як база даних, яка передбачає шість семантичних категорій: абстрактні відносини, простір, матерія, розум, воля, чуттєві та моральні сили. Усередині цих категорій будуються деревоподібні ієрархії підкласів. Для спрощення пошуку слова розташовуються в алфавітному порядку та забезпечуються інформацією про місце в ідеографічній класифікації.

Метрики й інструменти для аналізу деструктивного контенту в соціальних мережах та опис прикладного програмного забезпечення щодо виявлення деструктивного вмісту в постах представлено у [18]. Були зроблені такі кроки: підключення словників, нормалізація слів шляхом стемінгу. Однак основними параметрами у визначенні деструктивності контенту є відгуки користувачів мережі.

Відомі також методи, засновані на пошуку лексичної тональності у тексті за попередньо складеними тональними словниками та правилами з використанням лінгвістичного аналізу, де



спочатку визначається емотивна лексика, а потім текст оцінюється сукупно (наприклад, підсумовуванням значення тональності окремих речень) на основі кількісної шкали, градуйованої від негативної до позитивної оцінки [19]. Для аналізу використовують списки правил, регулярні вирази та спеціальні правила зв'язування тональної лексики в реченні.

Підвищення точності оцінювання тональності тексту можна досягти шляхом введення ваг для термінів. Найбільшу вагу повинні мати терміни, які краще відповідають тематиці тексту [6].

Аналіз показує, що автоматичні комплекси для визначення тональності мають такі недоліки:

- великий обсяг навчальних даних;
- складність навчання та неоднозначність результату навчання: з одного боку, збільшення обсягу навчальних даних дозволяє точніше налаштувати комплекс, з іншого боку, навчання залежить від ходу установки (викладача) і може призвести до різних результатів;
- навчання відбувається за алгоритмом "чорної скриньки", тобто його формалізація неможлива;
- складність складання семантичних словників.

Тому пропонується використовувати автоматизований комплекс оцінювання тональності, який працює за таким алгоритмом:

- 1) складання списку контенту;
- 2) формування бібліотек деструктивних конструкцій;
- 3) завантаження контенту та бібліотек конструкцій;
- 4) поділ контенту на об'єкти, розділені ключовим словом;
- 5) порівняння кожного об'єкта контенту з бібліотеками деструктивних конструкцій;
- 6) згодом записування кількості деструктивних слів у кожному об'єкті контенту до змінної;
- 7) аналіз і відсоткове оцінювання тональності тексту за такими розділами:
 - антиукраїнська пропаганда;
 - пропаганда екстремізму та тероризму;
 - критика існуючого порядку у країні;
 - розпалювання міжнародної ворожнечі;
 - заклики до насильства та суїциду;
 - заклики до повалення влади;
 - підриг духовно-моральних цінностей;
 - пропаганда алкоголізму та наркоманії;
 - інші негативні прояви;
- 8) візуалізація результату.

Для текстового контенту аналіз базується насамперед на вивченні його структури, а саме наявності та частоти вживання слів, пов'язаних із різними темами деструктивного змісту (ДК) (екстремізм, націоналізм тощо). Важливу роль в

аналізі відіграє знакова організація тексту, яка, у свою чергу, включає три компоненти:

- синтактика, яка порівнює різні знакові системи часової послідовності;
- семантика, яка демонструє відношення семантичного значення з пропозиціональним змістом тексту;
- прагматика, яка порівнює текст із його комунікативними установками.

Наступним критерієм може бути визначення емоційного забарвлення тексту. Існують алгоритми автоматизованого оцінювання тональності тексту, в якій значення обчислюється в діапазоні від 0 до 100 %, але це призводить до великого розкиду значень. Тому зручніше використовувати оцінку параметра настрою в діапазоні від 0 до 1, де 1 – позитивна оцінка, тобто зміст не є деструктивним, а 0 – негативна (деструктивна).

Точність і якість системи аналізу тональності тексту оцінюється тим, наскільки добре вона узгоджується з думкою людини щодо емоційного оцінювання досліджуваного тексту. Для цього можуть використовуватися такі показники, як точність і повнота.

Формула для знаходження повноти:

$$Entirety = \frac{CEO}{TNO}, \quad (3)$$

де CEO (*correctly extracted opinions*) – правильно розпізнані думки, TNO (*total number of opinions*) – загальна кількість думок (як знайдених системою, так і не знайдених).

Точність обчислюється за формулою

$$Veracity = \frac{CEO}{TNOFS}, \quad (4)$$

де CEO (*correctly extracted opinions*) – правильно розпізнані думки, TNOFS (*total number of opinions found by system*) – загальна кількість думок, знайдених системою.

Таким чином, точність виражає кількість досліджуваних текстів, речень або документів, в оцінюванні яких думка системи аналізу тональності збіглася з думкою експерта. Причому, згідно з дослідженнями, експерти зазвичай погоджуються в оцінках тональності конкретного тексту в 78,4 % випадків. Тобто програма, яка визначає тональність тексту з точністю 70 %, робить це майже так само добре, як і людина.

У розгляді публікації, представленої у формі/або зображенні, необхідно враховувати параметри яскравості, контрасту, а також колірних поєднань, оскільки ці показники потенційно можуть використовуватися розповсюджувачами деструктивного контенту для привернення уваги мережі, користувачів.



Для визначення яскравості зображення використовують вираз [20]

$$Y = 0,299R + 0,578G + 0,114B. \quad (5)$$

Яскравість зображення виражають як середню яскравість усіх пікселів. Яскравість одного пікселя розраховують за формулою

$$Y = R_p + G_p + B_p, \quad (6)$$

де Y_p – яскравість пікселя; R_p – червоний колір; G_p – зелений колір; B_p – синій колір; p – індекс числа пікселів.

Тоді фізичну яскравість зображення (видимого), що містить N пікселів, можна обчислити за формулою

$$Y = \frac{1}{N} \sum_{p=1}^N 0,299R_p + 0,578 G_p + 0,114B_p, \quad (7)$$

де Y – яскравість зображення; p – індекс підсумовування; N – кількість пікселів зображення.

Для оцінювання контрастності можна використовувати дисперсію яскравості пікселів зображення:

$$\sigma^2 = \frac{1}{N} \sum_{p=1}^N (Y_p - Y)^2, \quad (8)$$

де σ – стандартне відхилення.

Параметри для оцінювання яскравості та контрастності змінюються на інтервалі від [0; 1].

Відношення стандартного відхилення σ до максимально можливого значення яскравості Y_{\max} дозволить виміряти контраст яскравості C :

$$C = \frac{2\sqrt{N} \sum_{p=1}^N (Y_p - Y)}{\sum_{p_{\max}=1}^N 0,299R_{p_{\max}} + 0,578 G_{p_{\max}} + 0,114B_{p_{\max}}}, \quad (9)$$

де C – контраст яскравості зображення; $R_{p_{\max}}$, $G_{p_{\max}}$, $B_{p_{\max}}$ – значення R_p , G_p , B_p , що відповідають максимальному значенню яскравості зображення Y_{\max} .

Критерій (9) є інтегральним і змінюється в межах від 0 до 1, де 0 – одноколірне зображення, а 1 – найбільш насичене.

Однак, аналізуючи вміст фотоконтенту щодо його помітності, має сенс розглядати значення контрасту яскравості (C) лише в діапазоні від 0,5 до 1, який характеризує зображення, бо завдяки своїй яскравості та контрастності привертають більше уваги користувачів порівняно із зображенням зі значенням параметра від 0 до 0,5.

Згідно з науковими дослідженнями в галузі психології, кольори можуть впливати на сприйняття інформації людиною [21, 22]. З погляду деструктивного впливу на користувача соціальної мережі, особливу увагу слід звернути на такі кольори:

- чорний – деструктивність, придушення, депресія, порожнеча, застосування сили як прояв слабкості й егоїзму;

- червоний – фізичне насильство, нетерпимість, хтивість;

- синій – ідеалізм, фанатизм, субординація.

Основним інструментом поєднання кольорів є коло кольорів [23], сконструйоване таким чином, що поєднання будь-яких кольорів, обраних із нього, добре виглядатиме разом. Головне, правильно використовувати поєднання кольорів.

Відповідно до теорії кольору гармонійні поєднання кольорів виходять:

- із будь-яких двох кольорів, розташованих один навпроти одного на колірному колі;

- у разі використання будь-яких трьох кольорів, рівномірно розподілених на колірному колі з утворенням трикутника;

- у випадку використання будь-яких чотирьох кольорів, що утворюють прямокутник.

Гармонійні поєднання кольорів називають колірними схемами. Ці схеми залишаються гармонійними незалежно від кута повороту, що дозволяє розширити можливі комбінації маніпуляцій свідомістю.

Якщо необхідно на чомусь зосередити увагу користувача, вони використовують техніку з компліментарними або додатковими кольорами. Це будь-які два кольори, один проти одного розташовані на колірному колі. Наприклад, синій і помаранчевий, червоний і зелений. Ці кольори створюють високий контраст, тому їх використовують, коли ви хочете щось виділитися. В ідеалі один колір застосовують як фон, а інший – як акцент.

Класична тріада – це поєднання трьох кольорів, які рівновіддалені один від одного на колірному колі. Наприклад, червоний, жовтий і синій. Схема тріади також має високу контрастність, але більш збалансована, ніж додаткові кольори. Принцип тут полягає в тому, що один колір домінує і підкреслюється на фоні двох інших. Така композиція виглядає живою навіть під час використання блідих і ненасичених кольорів.

Крім того, фотоматеріали можуть містити різні праворадикальні знаки та символи, які стосуються пропаганди певної забороненої організації чи ідеології. Це може бути як відомий символ свастики або якихось рун, так і менш відомі знаки, які є символами заборонених екстремістських організацій. Крім того, це можуть бути кількісні символи, якими користується певна група людей. Усе вказане пов'язано з тим, що в багатьох країнах заборонено діяльність деяких неонацистських організацій, а в громадських місцях заборонено демонструвати символіку праворадикальних активістів. Щоб обійти ці заборони, екстремістські групи й окремі особи використовують шифри для передачі своїх повідомлень та уникнення судового переслідування.



Такий спосіб обману завжди можна зустріти на футбольних стадіонах, де вболівальники носять футболки або вивішують плакати з подібними шифрами. Цифри часто належать до порядкового значення букв в алфавіті. Наприклад, найчастіше вживають цифри 14, 18, 28 і 88 або їхню комбінацію, що є закодованим гаслом білих нацистів, яке часто використовують як привітання, і досить часто його зустрічаємо в інтернеті.

Аудіоконтент теж має особливості. Перш за все, варто зазначити, що люди можуть сприймати аудіоконтент на ходу, що дозволяє довготривало впливати на користувача.

Відомо, що низькочастотні звуки, близькі до нижньої межі зони чутності, можуть викликати дискомфорт, страх, занепокоєння і навіть паніку, якщо діють на людину досить довго і з необхідною інтенсивністю [24].

У таблиці наведено кілька прикладів впливу на людину звуку певних частот.

Таблиця

Залежність стану людини від несучої частоти

Значення несучої частоти, Гц	Вплив на людину
1–2	Серцебиття
41	Підвищення рівня метаболізму
62	Збільшення фізичної сили
131, 147, 165–169	Важка депресія
170–185	Дискомфорт
196, 247	Комфортний стан
333	Натхнення
349, 698	Любов
440, 880	Уява
523	Бар'єр страху

Яскравим прикладом звукового впливу є відома гра, яка має різні назви: "Тихий дім", "Розбуди мене о 4:20", "Море китів", "U19", "F57", "Синій кит" та ін. Можливо, сама назва пов'язана зі співом кита. Частота звуку, який видає синій кит, становить від 8 до 20 Гц, він не чутний людським вухом, але викликає занепокоєння, паніку і страх. За даними ЗМІ, завдання з 30 по 49 у цій грі пов'язані саме з прослуховуванням звуків певної частоти та нанесенням собі фізичної травми різанням. У цьому контексті видається цілком очевидним, що на учасників гри – дітей – вплинули, змушуючи їх слухати низькочастотний спів китів.

Тексти пісень також можуть становити загрозу. Його слід аналізувати за методами, наведеними вище. Тут постає додаткове завдання розпізнавання слів у аудіоконтенті.

Оцінювання відеоконтенту є найскладнішим, оскільки відеоматеріали є поєднанням текстового контенту, аудіоконтенту та зображень, які динамічно змінюються. Тому він повинен не тільки об'єднати всі перераховані вище показники, але й врахувати швидкість зміни кадрів, тобто різкі переходи від одного зображення до іншого, що супроводжується різними візуальними ефектами.

Однак найважливішим критерієм для будь-якого типу контенту є наявність у публікації ознак забороненого матеріалу. Якщо система виявить подібність цього контенту з відомими негативними матеріалами, його необхідно негайно передати на контент-аналіз.

3. ВИСНОВКИ

Представлені в роботі ознаки первинного оцінювання текстового, аудіо- та відеоконтенту соціальних мереж, а головне, тональність, з якою в них подається інформація, дозволять знизити ступінь суб'єктивності в оцінюванні і тим самим зменшать помилки розпізнавання деструктивного змісту.

Наступним важливим кроком у боротьбі з деструктивним контентом є створення відповідного вітчизняного програмного забезпечення з використанням представлених методів і параметрів, яке має застосовуватися як інструмент для моніторингу.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Браїловський М. М., & Толюпа С. В. (2021). Проблеми аналізування та прогнозування інформаційно-психологічних впливів у соціальних мережах, міжнар. наук.-практ. конф. "Інформаційна безпека та інформаційні технології": матеріали конференції". Харків – Одеса: Харківський нац. екон. ун-т ім. Семена Кузнеця. 99–103.
- [2] Yue L., Chen W., Li X., Zuo W., & Yin M. (2018, Jul.). A survey of sentiment analysis in social media. *Knowl. Inf. Syst.*, vol. 60, 617–663.
- [3] Tedmori, S., & Awajan, A. (2019). Sentiment analysis main tasks and applications: A survey *J. Inf. Process. Syst.*, vol. 15, no. 3, pp. 500–519.
- [4] Stefano B. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining, *Proceedings of LREC :conference*, pp. 2200–2204.
- [5] Pang, B., & Lee, L. (2004). A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts // *Proceedings of the Association for Computational Linguistics (ACL)* : article, pp. 271–278.
- [6] Liu, B. (2010). Sentiment Analysis and Subjectivity. Eds.: N. Indurkha, F. J. Damerau. *Handbook of Natural Language Processing*, pp. 38.
- [7] Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2. 16–17.
- [8] SenticNet (2009). <http://sentic.net/about/>.
- [9] Cambria, E., Poria, S., Bajpai, R., & Schuller, B. (Dec. 2016). SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives. *In Proc. 26th Int. Conf. Comput. Linguistics, Tech. Papers (COLING)*. Osaka, Japan: The



COLING 2016 Organizing Committee, pp. 2666–2677. <https://www.aclweb.org/anthology/C16-1251>.

[10] Gatti, L., Guerini, M., & Turchi, M. (Oct. 2016). SentiWords: Deriving a highprecision and high coverage lexicon for sentiment analysis, *IEEE Trans. Affect. Comput.*, vol. 7, no. 4, 409–421.

[11] Гофайзен О. В., & Пилявський, В. В. (2012). Область кольорів, передаваних системами цифрового телебачення. *Цифрові технології*, 11, 47–70.

REFERENCES

[1] Brailovsky M. M., & Tolyupa S. V. (2021). Problems of analyzing and forecasting informational and psychological influences in social networks, International Scientific and Practical Conference [Information Security and Information Technologies]: conference materials. Kharkiv – Odesa: Semyon Kuznets Kharkiv National University of Economics [in Ukrainian].

[2] Yue L., Chen W., Li X., Zuo W., & Yin M. (2018, Jul.). A survey of sentiment analysis in social media. *Knowl. Inf. Syst.*, vol. 60, 617–663.

[3] Tedmori, S., & Awajan, A. (2019). Sentiment analysis main tasks and applications: A survey *J. Inf. Process. Syst.*, vol. 15, no. 3, pp. 500–519.

[4] Stefano B. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining, *Proceedings of LREC :conference*, pp. 2200–2204..

[5] Pang, B., & Lee, L. (2004). A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts // *Proceedings of the Association for Computational Linguistics (ACL)* : article, pp. 271–278.

[6] Liu, B. (2010). Sentiment Analysis and Subjectivity. Eds.: N. Indurkha, F. J. Damerau. *Handboor of Natural Language Processing*, pp. 38.

[7] Pang B., & Lee L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retreval*, 2. 16–17.

[8] SenticNet (2009). <http://sentic.net/about/>.

[9] Cambria, E., Poria, S., Bajpai, R., & Schuller, B. (Dec. 2016). SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives. In Proc. 26th Int. Conf. Comput. Linguistics, Tech. Papers (COLING). Osaka, Japan: The COLING 2016 Organizing Committee, pp. 2666–2677. <https://www.aclweb.org/anthology/C16-1251>.

[10] Gatti, L., Guerini, M., & Turchi, M. (Oct. 2016). SentiWords: Deriving a highprecision and high coverage lexicon for sentiment analysis, *IEEE Trans. Affect. Comput.*, vol. 7, no. 4, pp. 409–421.

[11] Hofeisen, O. V., & Pylyavskiy, V. V. (2012). The range of colors transmitted by digital television systems. *Digital technologies*, 11, 47–70 [in Ukrainian].

Стаття надійшла до редколегії

20.01.2023

Methods of destructive influence and content protection in social networks

The work is devoted to the topical issue of protection of society from destructive influences in social networks. It has been proven that in order to prevent such influences, it is necessary to evaluate not only the materials contained in texts, audio and video content, but also the tone of the information disseminated in them. Tonality is one of the characteristics of thought or emotion and is currently used as the main parameter for classifying texts by thematic categories. This is due to the convenience and efficiency of classification according to this principle and greatly simplifies other tasks for the classification of text posts. Formed categories of tonal assessment. The possibility of assessing mood by one-dimensional or multidimensional parameter is proved. The possibility of using two approaches to classify the tonality of the text is shown. The first is the automatic analysis of digital texts by computers with elements of machine learning or teacher training. The second - with the help of dictionaries (thesauri) connected to the program. The shortcomings that occur in automatic systems for determining the key are given. A new algorithm for automatic complexes is proposed. It is noted that if the quality of the text analysis system is assessed by how well it agrees with a person's opinion on the emotional assessment of the studied text, then such indicators as accuracy and completeness can be used. Calculations of these parameters are given. When considering a publication presented in the form of an image, it is necessary to take into account the parameters of brightness, contrast, and color combinations, as these indicators can potentially be used by distributors of destructive content to attract the attention of the network. users. It is noted that the photo materials may contain various radical signs and symbols that may relate to the propaganda of a particular prohibited organization or ideology. The analysis of possible methods of influence on consciousness of the user by social networks is carried out.

Keywords: destructive effects of audio and video content; social networks; analysis of the tonality of the text; methods of analysis.



Микола Браїловський,
канд. техн. наук, доц.
Доцент кафедри кібербезпеки та захисту інформації факультету інформаційних технологій Київського національного університету імені Тараса Шевченка. Київ, Україна.

Mykola Brailovsky,
PhD (Engin.), Associate Prof.
Associate Professor of the Department of Cyber Security and Information Protection, Faculty of Information Technologies, Taras Shevchenko Kyiv National University. Kyiv, Ukraine.



Володимир Хорошко,
д-р техн. наук, проф.
Професор кафедри безпеки інформаційних технологій Національного авіаційного університету. Київ, Україна.

Volodymyr Horoshko,
Dr. Sci. (Engin.), Prof.
Professor of the Information Technology Security Department of the National Aviation University. Kyiv, Ukraine.