

УДК 004.912

Т.М. Заболотня<sup>1</sup>, А.Ю. Бартков'як<sup>2</sup>, А.В. Соколовська<sup>3</sup><sup>1</sup> НТУУ «КПІ», м. Київ, Україна, tatiana104@yandex.ua;<sup>2</sup> НТУУ «КПІ», м. Київ, Україна, andriy.bartkovyayak@hotmail.com<sup>3</sup> НТУУ «КПІ», м. Київ, Україна, anetkanfetka@gmail.com

## СПОСІБ ВИЗНАЧЕННЯ КОМПЛЕКСНОЇ ОЦІНКИ ТОНАЛЬНОСТІ ВІДГУКІВ ІНТЕРНЕТ-КОРИСТУВАЧІВ

Проведено аналіз специфіки Інтернет-коментарів, запропоновано спосіб оцінювання тональності відгуків, що базується на використанні словників «гостропозитивних» та «гостронегативних» слів, залученні додаткової фільтрації вхідного набору коментарів, а також на врахуванні додаткових параметрів відгуків для зниження впливу нерелевантних коментарів на загальну оцінку тональності.

АНАЛІЗ ТОНАЛЬНОСТІ ТЕКСТУ, ЕМОЦІЙНЕ ЗАБАРВЛЕННЯ ТЕКСТОВИХ ДАНИХ, КОМЕНТАР, ВІДГУК

### Вступ

Поточний стан інформатизації сучасного суспільства характеризується, в першу чергу, масовим залученням Інтернет до всіх сфер діяльності людини. Можливості всевітньої мережі широко використовуються для роботи та розваг, для шопінгу, бронювання та покупки квитків, спілкування, отримання інформації про новини, установи, факхівців чи події, які нас цікавлять, тощо.

Важливим інструментом, який не дозволяє загубитися в цьому морі інформації, є коментарі. Зазвичай, відвідувачі сайтів мають можливість прокоментувати новини, блоги, фото, фільми, товари Інтернет-магазинів, а також залишити відгуки про різноманітні заклади (ресторани, салони краси), якість роботи конкретних спеціалістів (наприклад, лікарів), враження від певної події. Це дає змогу людям обмінюватися думками щодо предмету коментування, висловлювати своє ставлення до нього та дізнаватися погляди інших. Також коментарі широко використовуються різними компаніями для отримання «зворотного зв'язку» (англ. «feedback») від користувачів щодо своїх продуктів чи послуг, щоб в подальшому врахувати їх побажання та зауваження і поліпшити свій сервіс.

Проте в останні роки стрімке зростання чисельності Інтернет-аудиторії та відповідне збільшення кількості відгуків відвідувачів сайтів відчутно ускладнило задачу модераторів коментарів, призвело до збільшення фінансових та часових витрат на їх обробку, а часом і взагалі унеможливило ручний аналіз вмісту коментарів. Таким чином, розроблення методів, алгоритмів, а також програмних засобів автоматизованої обробки відгуків Інтернет-користувачів є актуальною задачею на сьогоднішній день.

### 1. Постановка задачі

В більшості випадків відвідувачі сайтів ознайомлюються з коментарями для того, щоб дізнатися, яке загальне враження справив на інших людей предмет обговорення (фільм, ресторан, товар, програмний продукт тощо). Цим пояснюється тенденція до проведення оцінювання тональності відгуків користувачів в рамках автоматизованої обробки вмісту секції коментарів на сайтах. Наразі існують програмні сервіси для пошуку

згадувань заданої фрази в контексті позитивно або негативно забарвлених повідомлень в мережі Інтернет, проте дані сервіси дають дуже загальну оцінку емоційного забарвлення тексту (позитивну, негативну або нейтральну) та не беруть до уваги специфіку відгуків користувачів. Таким чином, з огляду на вищезазначене, метою даної роботи стало підвищення точності оцінювання тональності відгуків відвідувачів сайтів шляхом розроблення нового способу визначення оцінки емоційного забарвлення коментарів Інтернет-користувачів з урахуванням властивих їм рис.

У відповідності до поставленої мети **задачами дослідження** є:

- вивчення специфіки відгуків Інтернет-користувачів та формулювання переліку їх характеристик, які можуть бути враховані під час визначення їх тональності;
- дослідження особливостей коментування предметів та подій представниками різних груп користувачів, пошук шляхів використання цих даних під час оцінювання емоційної забарвленості відгуків;
- аналіз допоміжних інструментів та ресурсів, використання яких впливає на ефективність визначення тональності текстових даних;
- розроблення способу визначення тональності відгуків Інтернет-користувачів на основі комплексного урахування характеристик відгуків та специфіки висловлювання думок різними групами відвідувачів сайтів, а також дослідження ефективності використання запропонованого способу за критерієм точності оцінювання емоційного забарвлення текстових даних.

### 2. Характеристики відгуків Інтернет-користувачів

Відгуки Інтернет-користувачів про певні події, продукти чи послуги мають низку характеристик, які автори вважають за доцільне враховувати при оцінюванні емоційного забарвлення коментарів.

#### 1. Час, коли було залишено коментар.

Коментарі, котрі були залишені на сайтах одразу після виходу в світ будь-якого продукту чи після якоїсь значущої події, зазвичай характеризуються підвищеною емоційністю (як в позитивний, так і в негативний бік), що

погано впливає на коректність загальної оцінки тональності всієї сукупності відгуків.

### 2. Сфера діяльності, якої стосується коментар.

Автори вважають, що негативний відгук, в якому згадується заподіяння шкоди здоров'ю чи життю людини, повинен мати набагато більший «від'ємний» вплив на формування усередненої оцінки тональності коментарів, ніж негативний відгук про фільм, пісню чи кліп.

Крім того, в кожній сфері діяльності людини є свої специфічні ключові слова, вживання яких може надавати коментарю позитивної чи негативної забарвленості (наприклад, для косметології негативний відтінок відгуку може надати наявність в ньому таких слів як: опік, почервоніння, висип, запалення, інфекція і т.д.).

### 3. Анонімність відгуків.

Згідно опитування Pew center [1] у вересні 2013 року, чверть Інтернет-користувачів пише свої коментарі анонімно. Користувачі молодшого віку (до 30 років) виявляють більше небажання пов'язувати своє реальне ім'я з онлайн-висловлюванням (до 40%). Анонімність Інтернет-коментування призвела до появи феномену, який психолог Джон Салер назвав «ефектом мережевої розкутості». У момент, коли людина приховує свою ідентичність, вона також позбавляється звичних стримувальних факторів своєї поведінки, що може створювати культуру агресії та знущання і унеможлиблювати осмислену дискусію. Проте існує також і інша доведена точка зору, що анонімність заохочує користувачів до активної поведінки, оскільки в такий спосіб вони відчують спільну ідентичність і можуть не боятися, що виділяться з-поміж інших. Також анонімність може розвивати креативне мислення та спричинити ризикові вчинки користувачів [2].

Зважаючи на дані дослідження, автори вважають за доцільне додати можливість врахування анонімних відгуків (чи, навпаки, їх ігнорування) при визначенні загальної оцінки тональності коментарів Інтернет-користувачів щодо заданого предмету чи події.

### 3. Особливості коментування предметів та подій різними групами Інтернет-користувачів

Дослідники Вісконсинського університету в Медісоні виявили, що манера висловлювання думок користувачами в секції коментарів відчутно впливає на ввічливість та адекватність наступних відгуків відвідувачів сайтів (незалежно від їх анонімності): чим грубіше коментарі під статтею, тим сильніше читачі починають розходитися в оцінках її змісту. Цей феномен вчені назвали «ефектом грубості» [3]. В рамках даного дослідження розглянуто можливі прояви цього феномену на теренах Інтернету. Нижче наведені найбільш поширені явища, з якими щодня доводиться стикатися користувачам всесвітньої мережі і негативний вплив яких на коректність оцінки тональності відгуків відвідувачів сайтів є значущим.

**Тролінг** (дослівний переклад з англ. «trolling» означає «ловлю риби на блешню») – форма соціальної провокації в мережевій комунікації, використовувана як

персоніфікованими учасниками, зацікавленими в більшій впізнаваності, публічності, епатажі, так і анонімними користувачами без можливості їх ідентифікації.

**Флейм** (від англ. «flame» - вогонь) - «суперечка заради суперечки», обмін повідомленнями в місцях багатокористувацького мережевого спілкування (наприклад, Інтернет-форуми, чати, соціальні мережі), що представляє собою словесну війну; нерідко вже не має відношення до первісної причини спору. Повідомлення флейму можуть містити особисті образи, часто направлені на подальше розпалювання сварки. Іноді застосовується в контексті тролінгу, але частіше флейм спалахує просто через образу на віртуального співрозмовника.

**Флуд** (від неправильно вимовного англ. «flood» - потік) – не пов'язані з основною тематикою обговорення повідомлення в Інтернет-форумах і чатах, що часто характеризуються великими обсягами написаного тексту. Інколи флудом вважається кілька повідомлень підряд від однієї особи.

**Холівар** (англ. «holy war» - священна війна, крім того, релігійні війни) - загальна назва суперечок між людьми, котрі є прихильниками діаметрально протилежних думок, які вони не бажають змінювати. Такий спір принципово безглуздий, оскільки жоден з учасників дискусії не збирається вислуховувати і обдумувати аргументи свого опонента, а прагне тільки максимально привабливо виглядати в очах читачів.

Також до повідомлень, що можуть відчутно спотворювати оцінку тональності коментарів, відноситься **спам** (англ. «spam») – масова розсилка повідомлень рекламного чи іншого характеру людям, які не висловили бажання її одержувати.

Крім вищезгаданих явищ на даний момент можна виділити декілька груп користувачів, що своєю не завжди коректною поведінкою привертають до себе увагу, а відтак їх коментарі вирізняються неадекватними оцінками предметів чи подій, які обговорюються.

**Хейтери** (англ. «hater» - ненависник) – термін, який відноситься до групи людей, що висловлюють свою ненависть на форумах та в соціальних мережах. Хейтери відрізняються від тролів, які прагнуть привернути до себе увагу шляхом провокаційних коментарів, оскільки вони, як правило, мають дуже негативну думку щодо предмету обговорення і висловлюють свої погляди шляхом розміщення на сайтах негативних коментарів без врахування думки та аргументів опонентів, на яких вони здійснюють атаку. Коментарі, зазвичай, є надмірно емоційними, агресивними і/або жорстокими.

**Фанати** (фан, фен; скор. від «фанатик», лат. «fanaticus» - «одержимий, шалений», грец. «φανατισμός» - «сліпа віра») – люди, що мають особливо підвищений потяг до певного об'єкту (кумиру). Про таких кажуть, що він «фанатіє» від чого-небудь. Об'єктом потягу може бути відома особистість (або група), спортивні клуби, мистецтво (музика, живопис, кінематограф і т. д.), різні ідеї. Фактично, об'єктом потягу і поклоніння фанатів може бути все що завгодно.

Зазвичай, активність груп фанатів та хейтерів є більшою під час прем'єр (фільмів, альбомів музичних гуртів або співаків, серіалів тощо), випуску певних продуктів (наприклад презентація нового iPhone, нових колекцій відомих брендів одягу, косметики і т.д.) або відкриття якогось закладу (ресторану, салону краси, магазину та ін.).

Ще одним поняттям, що відноситься до галузі аналізу коментарів Інтернет-користувачів, є «хайп-трейн». *Хайп-трейн* (англ. «hype train» – «галасливий потяг») – сленговий термін, що використовується для опису високого рівня очікування майбутнього випуску або анонсу нового продукту. Частіше за все використовується по відношенню до фільмів, серіалів, комп'ютерних ігор тощо.

Зважаючи на описані фактори, необхідним є, по-перше, врахування деяких з них при визначенні оцінки тональності окремо взятих відгуків, а, по-друге, зниження впливу неадекватних та нерелевантних коментарів на формування усередненої оцінки Інтернет-спільнотою певного продукту чи події, адже діяльність спамерів, флудерів, фанатів, хейтерів і т.д. ускладнює модерацію коментарів та отримання об'єктивної оцінки продукту чи послуги.

#### 4. Спосіб визначення комплексної оцінки тональності Інтернет-коментарів

У даній роботі автори вважають за доцільне виділити три етапи визначення усередненої оцінки тональності відгуків Інтернет-користувачів:

- попереднє фільтрування заданого набору коментарів за заделегідь обраними параметрами;
- визначення тональності кожного відібраного відгуку за допомогою класичних алгоритмів оцінювання емоційної забарвленості текстових даних;
- формування комплексної оцінки тональності відгуків Інтернет-користувачів.

*Попереднє фільтрування* відгуків автори пропонують проводити за наступними характеристиками.

1. *Анонімність відгуку* – необхідно передбачити в алгоритмі визначення тональності коментарів можливість включати до переліку оброблюваних відгуків ті, автори котрих не вказали свого імені (або нікнейму). Це впливає на кінцеву усереднену оцінку тональності сукупності відгуків, оскільки, зазвичай, незареєстровані користувачі менш тактовно висловлюють свої думки.

2. *Довжина відгуку* – на думку авторів потрібно мати змогу виключати з переліку оброблюваних відгуків ті, що є коротшими за встановлену довжину тексту (адже написати декілька коротких та неаргументованих коментарів набагато простіше, ніж один виважений та обдуманий коментар). Це зменшує вплив на середню оцінку тональності відгуків записів, створених автоматичними генераторами коментарів (ботами), та спаму.

3. *Наявність кількох відгуків одного автора* – врахування даного параметру вимагає підтримки двох варіантів обробки відгуків одного автора:

- а) аналізувати всі відгуки незалежно від їх авторства;
- б) аналізувати лише перший коментар автора, відкинувши всі інші (скоріше за все, оглядач висловив

свою думку у першому коментарі, а в подальших – він просто вступає в дискусію з іншими користувачами, що може бути тролінгом або флеймом, флудом, спамом чи холліваром, крім того, у подальших коментарях може бути висловлене ставлення вже не до продукту, що коментується, а до інших оглядачів або їх думок).

*Визначення базової тональності* відфільтрованих на першому етапі відгуків пропонується здійснювати на основі класичних алгоритмів оцінювання емоційної забарвленості текстових даних.

Крім цього, на другому етапі автори пропонують впровадити кроки щодо врахування *сфери діяльності людини, до якої належить предмет обговорення* – дану специфіку можна враховувати за допомогою створення та використання користувацьких словників «гостро-позитивних» та «гостронегативних» слів та словосполучень, характерних для сфери діяльності, якої стосується коментар (наприклад, для сфери харчування: «отруєння», «біль у животі» та інші подібні слова несуть негативний зміст), а також визначення коефіцієнтів впливу кожного такого слова/словосполучення на оцінку емоційного забарвлення коментаря. Ці вагові коефіцієнти мають враховуватись під час визначення емоційної забарвленості тексту коментарів і впливати на результуючу оцінку. Для цього автори пропонують модифікувати класичні алгоритми оцінювання тональності тексту, в основі яких лежить машинне навчання з учителем.

Описані нижче характеристики відгуків можуть бути використані для *визначення комплексної*, більш точної, оцінки їх тональності.

1. *Час створення коментаря* – на думку авторів, чим ближче дата створення коментаря до дати виходу обговорюваного продукту чи значущої події, тим менш корисним буде вміст відгуку (наприклад, у перший тиждень після релізу продукту варто зменшити вплив коментарів на усереднену оцінку на 50%, у подальші три тижні – на 25%). Введення даного параметра дозволяє зменшити вплив думок фанів та хейтерів на формування середньої оцінки тональності відгуків.

2. *Наявність оцінки предмету чи події, що обговорюється, автором коментаря* – на багатьох сайтах є можливість, окрім написання коментаря, залишити оцінку (наприклад, 10/10). Врахування цієї оцінки при визначенні тональності відгуку дає змогу порівняти відповідність його тексту до оцінки та відкинути амбівалентні коментарі (наприклад, коли коментар є позитивним, а оцінка оглядача – низькою).

Таким чином, основною ідеєю способу оцінювання тональності відгуків Інтернет-користувачів, що пропонується, є комплексне врахування їх характеристик для попереднього відсіювання малозначущих відгуків, а також для корегування базової оцінки емоційної забарвленості кожного з них з метою зниження впливу неадекватних та нерелевантних коментарів на усереднене значення оцінки тональності сукупності відгуків щодо заданого предмету обговорення. Авторі допускають

можливість використання не всіх наведених вище характеристик при програмній реалізації запропонованого способу. В залежності від особливостей вмісту вхідного масиву коментарів, що аналізуються, емпіричним шляхом можна обрати свій набір фільтрів та коефіцієнтів, що забезпечить отримання найвищих показників ефективності роботи відповідного програмного забезпечення.

### 5. Дослідження результатів застосування розробленого способу

Для вивчення ефективності застосування запропонованого способу визначення оцінки тональності відгуків Інтернет-користувачів пропонується розглянути декілька відгуків на фільм «Темний лицар повертається», що вийшов у прокат в США 16 липня 2012 року, оскільки ця стрічка є яскравим прикладом ефекту «хайп-трейн» через високі очікування після фільму «Темний лицар», який здобув схвальні оцінки як глядачів, так і кінокритиків. Коментарі (взяті з сайту IMDb [4]):

– 22 липня 2012, оцінка 10/10. «Подивився фільм тільки вчора ввечері, і це було абсолютно круто. Фільм починається з трохи похмурого початку, показуючи розбитий образ Брюса Вейна. Проте протягом перших 15 хвилин починається епік. Бейн просто чудовий! Історія починає потроху прищвидшуватися з супроводом від Ганса Цимера, який робить світ ще похмурішим. В кінці першої половини фільму можна побачити морально та фізично знищеного Брюса, а люди в Готемі відчувають на собі терор від Бейна. Коли Брюс знаходить вихід з пастки і Темний лицар повстає для того, щоб врятувати людей Готема, ми розуміємо, що фінал буде просто епічним! Гордон-Левіт грає чудово і його роль теж досить вагома в фільмі. Звісно, це фільм, який варто побачити, і його точно можна назвати головним блокбастером року. Нолан точно створив найкращу трилогію Бетмена і чудовий фінал до неї. Трилогія виглядає так, ніби вона була запланована правильно з самого початку. Я сумніваюся, що коли-небудь перезапуск хоча б наблизиться до рівня цих фільмів. Деякі люди жалілися, що фільм йде 164 хвилини, але для мене цей час пролетів непомітно. Рекомендую подивитися цей фільм в кінотеатрі, щоб насолодитися спецефектами»;

– 22 липня 2012, оцінка 1/10. «Фільм був настільки довгий та нудний, що я почав замислюватись про те, щоб піти з кінотеатру. Сценарій був занадто передбачуваний, і акторська гра іноді зовсім не вражала. Енн Гетевей була жахливою. Її акторська гра зовсім не переконувала, а вбрання виглядало як дешевий хелловінський костюм. Бейн, головний злодій цього фільму, був вбитий за декілька секунд і всі про нього забули. Виглядало це ніби сценарист хотів, щоб фільм закінчився якомога швидше в цьому моменті. Акторка, яка грала Міранду, що виявилися дочкою Рас аль Гула, повинна отримати винагороду за найгіршу смерть на екрані. Здається, це була найгірша акторська гра, яку я бачив. Блейк, який в цьому фільмі виявився Робіном, отримав доступ до печери Бетмена, всієї його високотехнологічної амуніції, якою не зможе

користуватися, адже ніколи не отримував спеціальних тренувань. В цьому фільмі Робін – це звичайний поліцейський без бійцівських навичок, але Брюс все одно дає йому доступ до печери і їде відпочивати з Жінкою-кішкою»;

– 16 лютого 2013, оцінка 10/10. «Це мій перший огляд, тому я хочу зробити його простим. «Темний лицар повертається» - це найепічніше завершення до супергеройського кіно, яке я коли-небудь бачив. Фільм неспішно будує свою історію і заслуговує уважного перегляду. Цей фільм дає вам те, що ви очікували – чудовий фінал до найкращої трилогії в історії. В кінці, я хочу подякувати Кристоферу Нолану за зображення Бетмена таким чином і створення найкращої фільмової адаптації цього героя».

За допомогою класичних алгоритмів оцінювання тональності текстових даних перший коментар (далі коментар 1, 2 і 3, відповідно) оцінений, як позитивний з коефіцієнтом 0.9 (у діапазоні від 0 до 1, де 0 негативний, а 1 – найбільш позитивний коментар; при переводі даного коефіцієнту у 10-бальну шкалу отримаємо приблизно 9), коментар 2 – негативний з коефіцієнтом 0.1 (1 з 10), коментар 3 – позитивний, коефіцієнт – 0.8 (8 з 10). Середня оцінка «позитивності» відгуків – 6 балів з 10. Згідно ж з власними оцінками авторів відгуків отримуємо дещо вище середнє значення «позитивності» коментарів - 7 балів з 10.

Звісно, наведена вибірка не може вважатися репрезентативною, адже для точного аналізу потрібний набагато більший набір коментарів, проте навіть на даних коментарях можна показати ефективність запропонованого способу. Розглянемо вплив визначених в п.4 параметрів на кінцевий результат оцінювання тональності відгуків.

1. Попередня фільтрація відгуків:

а) для прикладу навмисне обрані неанонімі коментарі;

б) довжина кожного з коментарів є достатньою для передачі аргументованої точки зору щодо вражень від переглянутого фільму, тому збільшується ймовірність, що коментар був написаний реальною людиною (у разі коротких коментарів «Чудовий фільм! 10 з 10» ймовірність їх написання ботами є значно вищою);

в) відібрані для прикладу коментарі є першими коментарями своїх авторів, в яких вони не звертаються до інших користувачів і не змінюють тему обговорення.

Таким чином, наведені для прикладу коментарі не будуть відкинуті з процесу обробки на етапі попередньої фільтрації відгуків. Втім, у разі аналізу великої кількості коментарів, запропоновані фільтри зменшили б вплив нерелевантних відгуків на формування загальної оцінки тональності коментарів.

2. Оцінювання загальної тональності відгуків за допомогою класичних алгоритмів вже виконано раніше (6 балів з 10). Врахування додаткових вагових коефіцієнтів для «гостропозитивних» та «гостронегативних» слів, що належать відповідним користувацьким словникам в

даному прикладі не мало місця через те, що наразі таких словників не існує.

3. Для уточнення отриманих оцінок емоційної забарвленості відгуків проаналізуємо їх додаткові характеристики:

а) дата прем'єри фільму – 16 липня 2012 року. Користувачі 1 та 2 залишили свої відгуки 22 липня 2012 року, тобто вони переглянули фільм у тиждень прем'єри та написали відгуки відразу, оцінивши фільм на 10 і 1 бал з 10 відповідно. Дані факти можуть свідчити про те, що користувач 1 є явним фаном даної трилогії, або ж перебував під ефектом так званого «хайп трейну». Користувач 2 навпаки був незадоволений, на що може бути 3 причини: не виправдання очікувань, через культовість другого фільму трилогії; перебування під впливом «хайп трейну» (очікував більшого, але очікування не були виправдані); користувач є хейтером. У разі введення додаткового вагового коефіцієнту можна знизити вагу даних коментарів при формуванні середньої оцінки відгуків на фільм. Враховуючи дату написання коментарів, було б доцільно значно знизити їх вплив (наприклад, встановити вагу даних коментарів – 10%). Коментар 3 написаний 16 лютого 2013 (більше, ніж через півроку після прем'єри та через два з половиною місяці після релізу на DVD (28 листопада 2012), тому можна вважати, що даний коментар має на 100% впливати на комплексну оцінку;

б) врахування власних оцінок авторів коментарів і зіставлення їх з базовими оцінками, отриманими за допомогою алгоритмів визначення тональності текстових даних, для коментаря 1 10/10 та 9/10 дало б у результаті усереднену оцінку 9,5, для коментаря 2 – 1/10 та 1/10 – 1, коментаря 3 – 10/10 та 8/10 – 9.

В результаті застосування запропонованого способу комплексна оцінка позитивності відгуків сягнула 8,375:

$$\frac{(9,5 \times 0,1) + (1 \times 0,1) + 9}{0,1 + 0,1 + 1} = 8,375.$$

На сайті IMDb у місяці прокату рейтинг даного фільму тримався в околі 9,5 балів і мав тенденції до спаду з віддаленням від дати прем'єри у зв'язку зі зниженням активності фанів та хейтерів. Наразі, через 4 роки після прем'єри, рейтинг фільму складає 8,5 на сайті IMDb і має усталену сформовану оцінку, яка майже не змінюється останні 2 роки.

### Висновки

У статті запропоновано новий спосіб визначення комплексної оцінки тональності відгуків Інтернет-користувачів, що базується на врахуванні їх специфіки (анонімності, довжині, даті написання тощо). Авторами пропонується проведення попередньої фільтрації незначущих відгуків до застосування класичних алгоритмів оцінювання тональності, а також введення додаткових вагових коефіцієнтів, спрямоване на зниження впливу на загальну оцінку нерелевантних та неадекватних коментарів. Також новим в роботі є запровадження використання словників «гостропозитивних» та «гостронегативних» слів, складених для аналізу коментарів з конкретної галузі діяльності людини.

На базі проведеного дослідження розроблено програмну бібліотеку методів визначення емоційної забарвленості текстових даних, а також веб-сервіс для визначення комплексної оцінки тональності відгуків Інтернет-користувачів, які реалізують описаний вище спосіб. Перспективним напрямом подальших досліджень даного питання автори вважають створення словників «гостропозитивних» та «гостронегативних» слів для оцінювання відгуків з різноманітних сфер діяльності людини.

**Список літератури:** 1. Poynter – Режим доступу: <http://www.poynter.org/2013/25-of-people-have-posted-anonymous-comments-pew-finds/222912>. – 10.03.2016 2. Blau, I. Studying Invisibly: Media Naturalness and Learning / I. Blau, A. Caspi // *Evolutionary Psychology and Information Systems Research*. – 2010. – Volume 24 of the series Integrated Series in Information Systems. – Pp. 193-216. 3. Anderson, A.A. The “Nasty Effect:” Online Incivility and Risk Perceptions of Emerging Technologies / A.A. Anderson, D. Brossard, D.A. Scheufele, M.A. Xenos, P. Ladwig // *Journal of Computer-Mediated Communication*. – 2014. – Volume 19, Issue 3. – Pp. 373-387. 4. IMDb Database Statistics. – Режим доступу: <http://www.imdb.com/stats>. – 11.04.2016.

*Поступила в редакцію 11.05.2016.*

УДК 004.912

**Способ определения комплексной оценки тональности отзывов Интернет-пользователей** / Т.Н. Заболотня, А.Ю. Бартков'як, А.В. Соколовская // *Бионика интеллекта: научн.-техн. журнал*. – 2016. – № 1 (86). – С. 8-12.

В статье предложен способ оценивания тональности отзывов Интернет-пользователей, который, помимо определения оценки эмоциональной окраски с помощью классических алгоритмов, предусматривает комплексный учет: характеристик самих отзывов; принадлежности автора отзыва одной из специфических групп Интернет-пользователей; наличия в тексте комментария слов, содержащихся в словаре «остропозитивных» и «остронегативных» слов и словосочетаний. Предварительная фильтрация отзывов по их характеристикам, а также внедрение набора коэффициентов для уменьшения веса комментариев представителей определенных групп Интернет-пользователей позволили повысить точность оценки тональности отзывов. Исследована эффективность применения предложенного способа оценивания на наборе отзывов о фильме..

Библиогр.: 4 назв.

UDC 004.912

**The way of the complex sentiment analysis of the Internet users' reviews** / T. Zabolotnia, A. Bartkovyak, A. Sokolovska // *Bionica intellecta Sci. Mag.* – 2016. – № 1(86). – P. 8-12.

This paper proposes the way for determining the tonality of the Internet users' reviews, which not only evaluates the emotional tone of reviews, using classic algorithms, but also considers following characteristics: properties of reviews, reviewer being in one of distinct groups of Internet users, presence of “heavily positive” or “heavily negative” words, contained in the dictionaries, in the text of the review. Due to preliminary filtration of reviews based on their characteristics, introduction of weight coefficients in order to decrease the weight of reviews by members of certain groups of Internet users, the overall precision of review tonality analysis has increased. The effectiveness of proposed way has been researched on the data set of movie reviews.

Ref.: 4 items.