

Ю.О. Сєров, Р.Б. Кравець, А.М. Пелешишин
Національний університет "Львівська політехніка",
кафедра інформаційних систем та мереж

МЕТОДИ АНАЛІЗУ ЕФЕКТИВНОСТІ ВЕБ-ФОРУМІВ

© Сєров Ю.О., Кравець Р.Б., Пелешишин А.М., 2009

Розглянута актуальна проблема дослідження та розроблення методів аналізу ефективності Веб-форумів.

Ключові слова – Веб-спільнота, форум, Інтернет-форум, аналіз ефективності, методи аналізу ефективності.

Article considers actual problem of Web-forum efficiency analysis research and developing.

Keywords – Web-community, forum, Internet-forum, efficiency analysis methods.

Постановка проблеми

На сучасному етапі розвитку мережі Інтернет та глобальної інформаційної системи World Wide Web, особливо з переходом до концепції розвитку WWW – Веб 2.0, дуже важливу роль відіграють Веб-спільноти, котрі складаються зі звичайних користувачів мережі. Веб-форуми є однією з найпопулярніших та найефективніших у плані представлення інформації у Вебі формою організації Веб-спільнот.

Основними об'єктами Веб-форуму є учасники та інформаційне наповнення, яке вони створюють: дискусії та повідомлення. Учасником Веб-спільноти будемо вважати людину, яка відвідує сайт Веб-спільноти, читає чи публікує інформацію у вигляді дискусій та повідомлень. Інформаційне наповнення Веб-форуму – це ієрархія підфорумів та повідомлень, дерево з вузлами двох типів: підфоруми і дискусії.

Сьогодні більшість сучасних Веб-форумів функціонують на основі популярних програмних рішень – CMS(Content Management System – Система управління інформаційним наповненням (СУІН)), серед яких можна назвати vBulletin, Invision, phpBB, XMB.

Для кожного власника та адміністратора важливим є завдання моніторингу та аналізу стану та процесу життєдіяльності його Веб-форуму. Проте жодна з перелічених СУІН не забезпечує адміністратора якісними засобами моніторингу та аналізу. Здебільшого за допомогою СУІН адміністратор Веб-форуму може отримати лише дуже обмежений набір статистичних даних (кількість повідомлень за день, кількість учасників, які відвідали форум протягом доби, загальну кількість повідомлень та дискусій Веб-форуму, найактивніші дискусії, найатрактивніші дискусії, найактивнішого учасника за день, множину учасників, які сьогодні святкують день народження, тощо), але жодна СУІН не дає можливості спостерігати велику кількість важливих показників (позицію форуму, динаміку створення повідомлень та дискусій, динаміку зростання інформаційного наповнення Веб-форуму, найактивніші підфоруми, динаміку активності учасників, наповнення підфорумів, розподіл аудиторії учасників по підфорумах тощо).

Аналіз останніх досліджень

Попри велику кількість наявних на сьогодні Веб-спільнот на основі форумів, лише окремі є ефективними. Тому дослідження та аналіз критеріїв ефективності Веб-спільнот є важливим й актуальним завданням, яке дасть змогу розробити методи підвищення ефективності Веб-спільнот.

Основними критеріями ефективності Веб-форумів є: цитованість, відвідуваність, кількість зареєстрованих учасників, об'єм інформаційного наповнення, швидкість приросту інформаційного наповнення, швидкість приросту кількості учасників.

Цитованість – критерій, який свідчить про авторитетність Веб-форуму, первинність та актуальність інформаційного наповнення Веб-форуму.

Відвідуваність – цей критерій ефективності відображає розмір аудиторії спільноти. Розрізняють такі види відвідуваності: безпосередній прихід, з пошуковця, з інших сайтів, з поштового клієнта, з RSS-агрегатора.

Кількість зареєстрованих учасників вказує на розмір аудиторії сайта-спільноти. Учасники поділяються на класи (**Активіст, Модератор, Автор, Критик, Пустомеля, Читач**) на основі таких даних: кількість створених повідомлень, кількість створених дискусій, кількість повідомлень за одиницю часу, інтенсивність створення повідомлень, оцінка діяльності інших учасників.

Об'єм інформаційного наповнення відображає креативність учасників спільноти. Збільшення об'єму інформаційного наповнення сприяє зростанню цитованості (авторитетності) і покращанню позиції у запитах до пошукових машинах.

Швидкість приросту інформаційного наповнення свідчить про активність та креативність учасників спільноти. Цей критерій визначається на основі нових дискусій та повідомлень за проміжок часу.

Швидкість приросту кількості учасників свідчить про підвищення популярності Веб-форуму.

Наведені критерії ефективності є основою для створення методів підвищення ефективності функціонування Веб-спільнот на основі форумів у мережі Інтернет.

Усі Веб-спільноти, зокрема і форуми, функціонують на основі програмних комплексів – систем управління інформаційним наповненням (СУІН), які забезпечують функціонування спільнот. Загалом усі СУІН, на основі яких функціонують Веб-форуми, мають складну схему даних, яка складається з великої кількості таблиць. Однак значна частина цих таблиць є довідниками та виконує допоміжну функцію.

Розглянувши та проаналізувавши наявні СУІН Веб-форумів, відкинувши таблиці-довідники та інші другорядні сутності, виділимо основні об'єкти Веб-спільноти як інформаційної системи. Узагальнена схема даних, яка забезпечує необхідний набір сутностей для аналізу поведінки учасників Веб-спільнот, має вигляд, зображений на рис. 1.

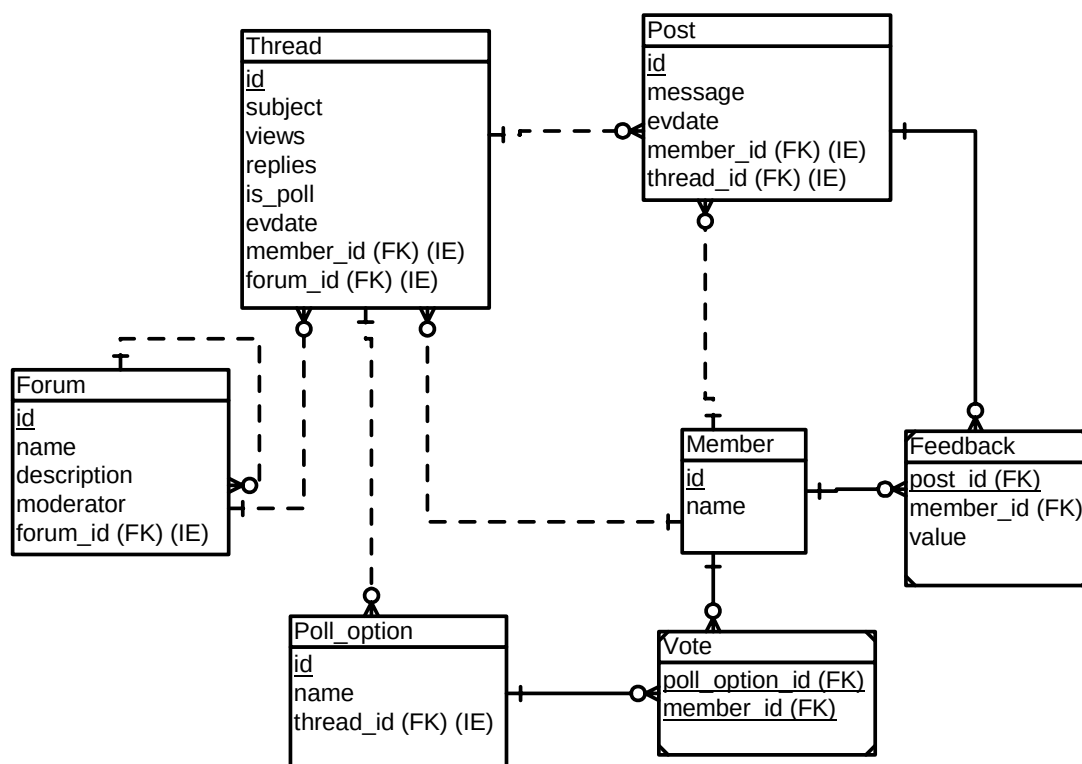


Рис. 1. Інформаційна модель Веб-форуму

Основними складовими такої інформаційної системи, як Веб-спільнота, є **Учасники (Members)** та **Інформаційне наповнення**.

Інформаційне наповнення структурно складається з Форумів (Forum), Дискусій (Thread) та Повідомлень (Post) [1].

У попередніх роботах авторів проаналізовано поведінку учасників Веб-спільнот, моделювання активності учасників Веб-спільноти за допомогою апарату нечітких множин, правила та методи класифікації учасників Веб-спільнот на основі Веб-форумів, а також наводяться приклади розроблення програмних комплексів для аналізу поведінки учасників спільнот [1–3, 8].

Серед усіх видів Веб-спільнот (чати, форуми, гостьові книги, блоги, соціальні мережі) найбільше досліджень стосується соціальних мереж. У багатьох роботах і зокрема в цих [5, 6] проводяться дослідження Веб-спільнот (у соціальних мережах) з метою визначення авторитету учасників цих Веб-спільнот та дослідження зв'язків між ними. Також важлива увага приділяється завданням візуалізації структури зв'язків між людьми у межах Веб-спільнот [7]. Проводити такі дослідження є можливість лише у Веб-спільнотах, де використовується механізм “зв'язку” між учасниками спільноти. Цей механізм характерний лише для соціальних мереж на зразок Facebook. Сенсом таких досліджень є виокремлення кращих у певному сенсі учасників спільноти з метою подальшого використання цих знань – необов'язково у межах Веб-спільноти.

Стосовно Веб-форумів ці дослідження рідко застосовні, оскільки соціальні мережі – це той тип Веб-спільнот, які орієнтовані на людей та особистісні зв'язки та відображають зв'язки людей у реальному житті, тоді як Веб-форуми об'єднують людей на основі інформації. Крім того, популярні СУІН не підтримують механізму створення зв'язків між учасниками Веб-форуму.

Для Веб-спільнот на основі Веб-форумів та їх власників (адміністраторів) важливішим є дослідження у напрямках дослідження життєдіяльності форуму, структури інформаційного наповнення та поведінки його учасників; моніторинг виникнення критичних та небажаних ситуацій, планування розвитку Веб-форуму, а також розроблення програмного забезпечення, яке б автоматизувало розв'язання цих задач. Сьогодні не розроблене програмне забезпечення, інтегроване у популярні СУІН, яке б розв'язувало ці задачі.

Цілі статті

Завдання цієї статті – описати основні методи аналізу ефективності Веб-спільнот та поведінки її учасників.

Аналіз ефективності Веб-форумів є складною комплексною задачею, яку розв'язують різними методами з урахуванням як внутрішньої, так і зовнішньої щодо форуму інформації. Розроблення методів аналізу ефективності Веб-форумів дасть змогу ефективно визначати їх стан, проблеми, які існують, та визначати рекомендовані напрями їх розвитку.

Отже, цілями цієї статті є:

- дослідження відомих методів аналізу ефективності Веб-форумів;
- розширення схеми даних для аналізу стану Веб-форуму;
- розроблення методів аналізу приросту інформаційного наповнення Веб-форуму;
- розроблення методів аналізу приросту учасників Веб-форуму;
- розроблення методів аналізу позиції Веб-форуму;
- апробація розробленого на основі даних Веб-спільноти “Форуму рідного міста”.

Основний матеріал. Аналіз стану Веб-форуму

Розширення схеми даних для аналізу стану Веб-форуму

Аналіз стану Веб-форуму зводиться до визначення його показників ефективності. Чотири з наведених критеріїв ефективності, а саме: кількість зареєстрованих учасників, об'єм інформаційного наповнення, швидкість приросту інформаційного наповнення, швидкість приросту кількості учасників, можна визначити на основі даних, що містяться в базі даних форуму, однак для цього треба розширити функціональність наявних СУІН. Розширення функціональності полягає у доопрацюванні програмної частини СУІН з метою розв'язання задач визначення ефективності

Веб-форуму. Розроблення нових програмних засобів, своєю чергою, передбачає розширення схеми бази даних для необхідного аналізу стану Веб-форуму, визначення показників ефективності Веб-форуму, а також моніторингу та можливого прогнозування його розвитку.

Розширена схема даних для аналізу стану Веб-форуму зображена на рис. 2.

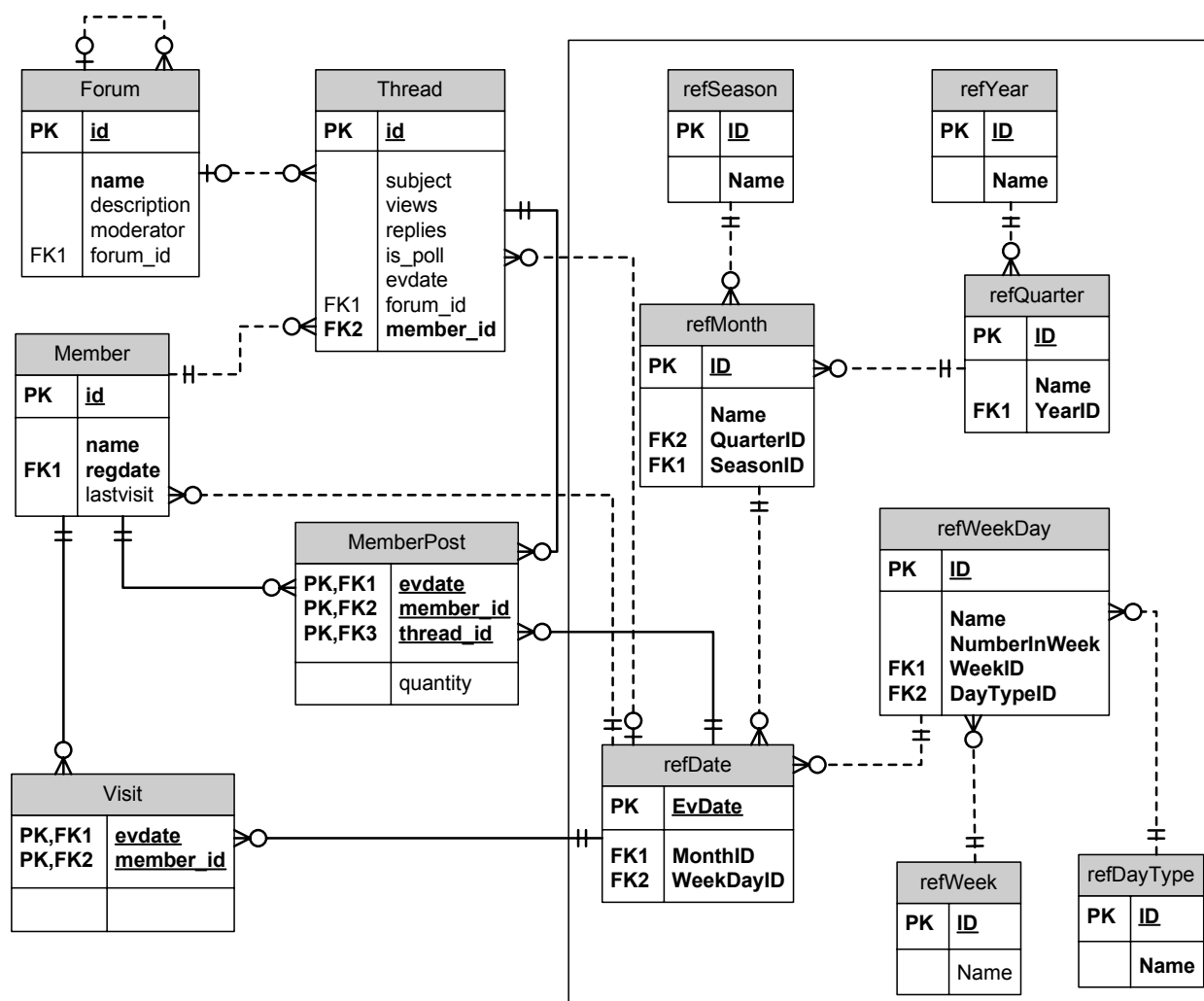


Рис. 2. Схема даних системи аналізу стану та процесу життєдіяльності Веб-форуму

Аналізувати стан Веб-форуму згідно з критеріями ефективності будемо у двох напрямках – окремо будемо аналізувати учасників Веб-форуму та його інформаційне наповнення. Своєю чергою, аналіз учасників Веб-форуму можна вести у двох напрямках – аналіз появи нових учасників та діяльність наявних. Аналізувати збільшення інформаційного наповнення будемо стосовно його поділу на типи, тобто зростання кількості повідомлень та збільшення кількості дискусій. Усі ці показники діяльності Веб-форуму є взаємозалежними. Скажімо, збільшення кількості учасників вказує на актуальність та атрактивність інформаційного наповнення Веб-форуму, і, своєю чергою, веде до збільшення кількості інформаційного наповнення. І навпаки, збільшення кількості цікавого контенту призводить до збільшення кількості відвідувачів, і як наслідок – учасників. Своєю чергою, зменшення приросту нових учасників призводить до зменшення активності наявних.

Для забезпечення стабільної діяльності та розвитку Веб-форуму необхідно, щоб показники ефективності на великих часових проміжках були неспадними.

Аналіз приросту учасників Веб-форуму та їх діяльності

Аналізуючи Веб-форум стосовно його учасників і їх діяльності, почнемо з аналізу приросту учасників. Першою дією людини, яку їй потрібно зробити для того, щоб стати учасником Веб-

форуму, є реєстрація. Переважно процедура реєстрації є типовою – необхідно задати своє ім'я та пароль, а також інші дані, які можуть бути обов'язковими або ні – такі як електронна пошта, дата народження тощо.

Аналіз приросту учасників

Досліджувати учасників Веб-форум почнемо з аналізу їх приросту. Як правило, одна людина дуже рідко може постійно бути активним учасником Веб-форуму, переважно з часом активність кожного учасника знижується. Тому для стабільної життєдіяльності та розвитку Веб-форуму необхідний постійний приріст учасників. Приріст учасників – показник, який демонструє популярність Веб-форуму, дає змогу власнику/адміністратору зрозуміти, чи правильно розвивається Веб-форум.

Приріст учасників є сенс аналізувати на середньо- і довгострокових часових проміжках – місяць, квартал, рік. Також потрібно аналізувати сезонність цього явища. З практичного погляду метою такого аналізу є дослідження динаміки зростання спільноти для прогнозування необхідної кількості модераторів, проведення рекламних кампаній та дослідження їх результатів.

a)

```
SELECT regdate,  
COUNT (*) FROM Member  
GROUP BY regdate
```

b)

```
SELECT  
    d.evdate,  
    COUNT (m.id) as member_number  
FROM Date as d  
    LEFT JOIN Member as m ON  
        m.regdate=d.evdate  
GROUP BY d.evdate
```

Запит для визначення кумулятивної кількості учасників:

```
SELECT  
    regdate,  
    COUNT(m1.regdate) ,  
FROM  
    Member as m1  
INNER JOIN Member as m2 ON  
    m1.regdate<=m2.regdate  
GROUP BY m1.regdate
```

Аналіз відвідуваності Веб-форуму

Наступним важливим показником стану Веб-форуму є його відвідуваність. Відвідуваність є важливим показником, який вказує на міру популярності Веб-форуму, актуальність його інформаційного наповнення та його унікальність.

Усі відвідувачі Веб-форуму поділяються на зареєстрованих і незареєстрованих. Аналізувати відвідуваність необхідно для цих двох типів відвідувачів.

- Незареєстровані учасники Веб-форуму потрапляють на нього, як правило, за допомогою пошукових серверів, прямих лінків з інших сайтів, з поштових розсилок, з рекламних джерел (не обов'язково Веб-ресурсів). Такий тип відвідуваності відповідно свідчить про унікальність та актуальність інформаційного наповнення Веб-форуму, зацікавленість інших людей людьми, чи інформацією, що міститься на форумі, якість поштових розсилок та інших видів реклами Веб-форуму.
- Зареєстровані відвідувачі Веб-форуму, тобто його учасники, відвідують форум з певною періодичністю. Цей вид відвідуваності, зокрема її період (частота), свідчить про прив'язаність учасника до Веб-форуму, період оновлення інформації, яка зацікавила учасника. Така інформація дає змогу визначати кількість постійних учасників Веб-форуму,

тобто таких учасників, середній період яких між відвідуваннями Веб-форуму не перевищує певного заданого часового інтервалу, наприклад, день, тиждень, місяць. Також опосередковано ця інформація дає змогу проаналізувати швидкість оновлення інформаційного наповнення Веб-форуму (появу нових повідомлень у дискусіях та виникнення нових дискусій).

На основі аналізу такого показника, як відвідуваність, визначатиметься кількість постійних відвідувачів форуму, як із зареєстрованих, так і із незареєстрованих учасників форуму, що дасть змогу визначити аудиторію форуму. При визначенні множини постійних учасників форуму потрібно визначити певний часовий проміжок, невідвідання форуму протягом якого виключатиме його з кола постійних учасників.

Крім того, можна визначати такий показник, як період відвідуваності з метою виявлення часу наступного зацікавлення учасників форумом, що на практиці означатиме період оновлення інформаційного наповнення форуму і міру “прив’язаності” учасника до форуму.

Запит для визначення відвідуваності за днями виглядає так:

```
SELECT
    evdate,
    COUNT(*)
FROM Visit
GROUP BY evdate;
```

Запит для визначення відвідуваності за місяцями виглядає так:

```
SELECT
    m.name,
    COUNT(*)
FROM Visit as v
    INNER JOIN Date as d ON d.evdate=v.evdate
    INNER JOIN Month as m ON m.id=d.MonthID
WHERE (v.evdate BETWEEN 'd1' AND 'd2')
GROUP BY (m.name);
```

Аналіз важливості учасника та визначення “ядра” форуму

Кожен учасник Веб-форуму тією чи іншою мірою корисний для спільноти, причому його корисність може бути як додатною, так і від’ємною. Корисність учасника може визначатися як об’єктивно, на основі дослідження його поведінки, інформаційного наповнення, яке він створив та класифікації [2], так і суб’єктивно, на основі оцінок інших учасників Веб-форуму та експертних оцінок.

Такий показник, як корисність, дає змогу ранжувати учасників Веб-форуму, з подальшим застосуванням цієї інформації для вживання адміністративних дій. Наприклад, найкорисніших учасників потрібно залучати до модерування форуму, стимулювати матеріально, водночас учасників з від’ємною корисністю потрібно видаляти з Веб-спільноти.

У найпростішому випадку корисність учасника Веб-форуму визначатимемо, аналізуючи кількість його повідомлень. Якщо вважати, що за принципом Парето [9] вісімдесят відсотків усього внеску в створення усього інформаційного наповнення створюють двадцять відсотків учасників Веб-форуму, то визначення ядра форуму – визначення групи учасників, що роблять найбільший внесок у формування інформаційного наповнення спільноти. Однак співвідношення 80-20 не є обов’язковим, можна просто визначати групу учасників, які роблять максимальний внесок у створення інформаційного наповнення.

Ранжування учасників за корисністю і визначення ядра форуму дає змогу адміністратору постійно знати список учасників, що приносять Веб-форуму максимальну користь. Така інформація необхідна і дуже корисна для управління Веб-форумом, оскільки Веб-спільнота тим і відрізняється від звичайного сайту, що адміністратор повинен враховувати думки та інтереси учасників. І щоб приймати корисні для спільноти рішення, треба дослухатися до тих учасників, які володіють більшим авторитетом, приносять більше користі для спільноти і роблять більший внесок у її розвиток.

Визначення “ядра” форуму реалізоване за допомогою запиту, який спирається на table_valued функцію.

Функція має такий вигляд:

```
(SELECT
  member_id,
  SUM(quantity) as post_quantity
FROM MemberPost
[WHERE evdate BETWEEN d1 AND d2
GROUP BY member_id] as MP
```

Запит для визначення “ядра” форуму виглядатиме так:

```
SELECT
  MP1.member_id,
  MP1.post_quantity
FROM MP(d1, d2)
  INNER JOIN MP(d1, d2) as MP2 ON
    MP2.post_quantity<=M1.post_quantity
GROUP BY
  MP1.member_id,
  MP1.post_quantity
HAVING SUM(MP2. post_quantity)<=core_threshold*
      (SELECT SUM (post_quantity)
        FROM MP(d1, d2))
```

Аналіз приросту інформаційного наповнення Веб-форуму

Інформаційне наповнення Веб-форуму складається з опитувань, дискусій та повідомлень, які створюють учасники спільноти. Приріст інформаційного наповнення є важливим критерієм ефективності та свідчить як про активність та креативність учасників, так і про актуальність тем, які обговорюються на форумі.

Аналіз приросту кількості повідомлень

Інтенсивність приросту кількості повідомлень свідчить про актуальність наявних на форумі дискусій та активність зареєстрованих учасників. Аналізувати приріст кількості повідомлень можна:

- для визначення підфорумів дискусій, у яких найбільше дискутують з метою визначення тематики, яка найбільше цікавить учасників Веб-форуму;
- для класифікації учасників за інтересами, дослідження сфери їхніх інтересів;
- для визначення активних учасників Веб-форуму з метою їх стимулювання;
- для виявлення конфліктних ситуацій.

Відсутність нових повідомлень може свідчити про відсутність актуальних та цікавих на певний момент часу дискусій, про малу кількість активних учасників, і навіть про технічні проблеми з доступом до форуму. Наявність дискусій без відповідей свідчить про відсутність серед учасників форуму цільової аудиторії для дискусій з такою тематикою і навпаки, наявність багатьох дискусій з великою кількістю повідомлень свідчить про відповідність тематики цих дискусій інтересам учасників форуму.

Запит для визначення приросту кількості нових повідомлень виглядає так:

```
SELECT
  d.evdate,
  SUM(MP1.quantity) as growth,
  SUM(MP2.quantity) as cumulative
FROM Date as d
  LEFT JOIN MemberPost as MP1 ON MP1.evdate=d.evdate
    LEFT JOIN MemberPost as MP2 ON
      MP2.evdate<=MP1.evdate
WHERE MP1.evdate BETWEEN d1 AND d2
GROUP BY d.evdate
```

Приріст кількості повідомлень слід аналізувати як на великих часових проміжках, так і щоденно.

Аналіз приросту кількості дискусій

Інтенсивність приросту кількості дискусій свідчить про креативність учасників Веб-форуму. Аналіз приросту кількості дискусій, як і аналіз приросту кількості повідомлень, дає змогу визначати предметні області, які викликають на певний момент часу найбільше зацікавлення серед учасників форуму.

Запит для визначення приросту кількості нових дискусій виглядає так:

```
SELECT
    d.evdate,
    COUNT (T1.*) as T_growth,
    COUNT (T2.*) as T_cumulative
FROM Date as d
    LEFT JOIN Thread as T1
    ON T1.evdate=d.evdate
    LEFT JOIN Thread as T2
    ON T2.evdate<=T1.evdate
WHERE d.evdate BETWEEN d1 AND d2
GROUP BY d.evdate
```

Аналіз приросту кількості дискусій слід здійснювати на більших часових проміжках, хоча для великих форумів такий вид аналізу є сенс виконувати і поденно.

Ці два види аналізу контенту форуму дають змогу приймати рішення про реструктурування дерева інформаційного наповнення форуму з метою його кращого подання та швидшого пошуку.

Аналіз позиції форуму

Застосувавши методику оцінки позиції сайту [10], опишемо метод визначення позиції форуму у WWW. Позиція форуму у WWW виглядає так:

$Pos(Forum) = (Th(Forum), Db(Forum), Aud(Forum), Auth(Forum), Rank(Forum, Th))$, де $Th(Forum)$ – тематика форуму або його регіональна приналежність; $Db(Forum)$ – інформативність форуму; $Aud(Forum)$ – аудиторія форуму; $Auth(Forum)$ – авторитетність форуму, $Rank(Forum, Th)$ – ранг форуму серед форумів з певної тематики чи тих, що належать до однієї територіальної одиниці.

Інформативність форуму

Оскільки сторінки з дискусіями Веб-форумів мають типову структуру, інформативність форуму $Db(Forum)$ залежатиме лише від:

- обсягів інформаційного наповнення;
- унікальності інформаційного наповнення;
- актуальності та новизни інформації;
- мультимедійного інформаційного наповнення.

Інформативність форуму $Db(Forum)$, з погляду пошукових машин, визначається як множина усіх сторінок усіх дискусій форуму, на яких є хоча б одне повідомлення без урахування цитованої інформації.

Кількісна оцінка об'єму інформаційного наповнення має такий вигляд:

$$\|Db(Forum)\| = \sum_{Pg_j^i \in Db(Forum)} \frac{Size(Pg_j^i) - Size(Quote(Pg_j^i))}{Block},$$

де Pg_j^i – j -та сторінка i -ї дискусії форуму, $Size(Pg_j^i)$ – фізичний розмір j -а сторінки i -ї дискусії (у байтах) без урахування елементів структурного та графічного форматування конкретної системи управління інформаційним наповненням форуму; $Size(Quote(Pg_j^i))$ – фізичний розмір цитат, зроблених учасниками Веб-форуму; $Block$ – розмір сторінки СУІН, на основі якої функціонує форум (у байтах).

Аудиторія форуму

Аудиторія форуму $Aud(Forum)$ визначається як сумарна кількість учасників форуму та постійних відвідувачів форуму, які не є зареєстрованими.

$$Aud(Forum) = \{Member(Forum) \cup Guest(Forum)\}.$$

Тоді як задача підрахунку кількості учасників є нескладною, а задача їх класифікації для визначення тих, кого можна зарахувати до аудиторії форуму, уже розв'язана [2], то задача підрахунку незареєстрованих учасників та їх класифікації є складнішою. Усіх незареєстрованих відвідувачів форуму умовно можна поділити на дві групи: ті, що відвідують форум регулярно і ті, що потрапили на нього одноразово (з пошуковців чи інших зовнішніх посилань). Розділити ці два типи гостей з метою визначення постійних відвідувачів можна на основі моніторингу та аналізу статистики на певному часовому проміжку відвідувань форуму (за ір-адресами). Для цього потрібно відняти від загальної кількості постійних відвідувачів кількість зареєстрованих учасників, що постійно відвідують форум.

$$Guest(Forum) = \{AllVisitors(Forum) - Member(Forum)\}.$$

Авторитетність форуму

Авторитетність форуму є агрегованою авторитетністю сторінок його дискусій. Алгоритми на зразок алгоритму PageRank [11] дають можливість оцінити авторитетність окремо взятої сторінки, але не форуму загалом. Крім того, при визначенні авторитетності певного сайту такі алгоритми спираються на авторитетність інших сайтів, які містять посилання на цей сайт. З Веб-форумами ситуація дещо інша. Оскільки Веб-форуми – це середовище живого спілкування людей, то і їх авторитетність слід визначати на основі їх згадування людьми на сторінках інших Веб-спільнот. Але оскільки сторінки Веб-спільнот, що містять посилання на сторінки Веб-форуму, як правило, мають низький ранг у пошукових сервісах, то спиратися потрібно на кількість переглядів цих сторінок.

Тому авторитетність форуму визначимо як сукупну авторитетність його сторінок:

$$Auth(Forum) = \{ \{Pg_j^i, Auth(Pg_j^i)\} \}, \text{ де } Pg_j^i - j\text{-та сторінка } i\text{-ї дискусії форуму, а } Auth(Pg_j^i) -$$

її авторитетність. Авторитетність кожної сторінки форуму визначатиметься як сумарна кількість посилань на неї з інших сайтів, помножених на їх авторитетність:

$$Auth(Pg_j^i) = \sum_{Pg_j^i \in Db(Forum)} Auth(Link(Pg_j^i)), \text{ причому авторитетність посилань з форумів, блогів та}$$

інших сайтів-спільнот визначається на основі кількості переглядів сторінок з посиланнями на сторінку Pg_j^i , а авторитетність усіх інших посилань визначається на підставі показників авторитетності, встановлених для них пошуковими системами (наприклад, PageRank).

Ранг форуму

Ранг форуму, який має конкретну тематику чи регіональну прив'язку, визначається як множина порядкових величин, кожна з яких визначає місце сайту серед конкурентів за певною ознакою: кількість зареєстрованих учасників, кількість загальних переглядів усіх дискусій форуму, загальна кількість повідомлень на форумі тощо.

$$Rank(Forum, Th) = \sum_{j=1}^{N(Opt)} \alpha(Opt_j) * Rank(Forum, Th, Opt_j), \text{ де } Opt_j - j\text{-а ознака, а}$$

$\alpha(Opt_j) \in [0;1]$ – ваговий коефіцієнт ознаки Opt_j . Рангом форуму буде його позиція у результатах пошукового запиту за тематикою після сортування за однією з ознак (кількість зареєстрованих учасників, кількість загальних переглядів усіх дискусій форуму, загальна кількість повідомлень на форумі) або усереднена позиція форуму після сортування за кожною з ознак.

Висновки

Сьогодні в мережі Інтернет існує величезна кількість Веб-форумів, тому дослідження та розроблення методів аналізу ефективності Веб-форумів є важливим й актуальним завданням. Розроблення методів аналізу дасть змогу власникам Веб-спільнот ґрунтовніше аналізувати стан Веб-форуму, швидше реагувати на виникнення критичних ситуацій, прогнозувати темпи розвитку. У статті наведено розширену схему даних Веб-форуму, необхідну для аналізу приросту учасників та інформаційного наповнення Веб-форумів, на основі існуючих досліджень розроблено методику визначення позиції форуму.

1. Kravets R. Typical ways of web-communities development / Kravets. R., Peleschyshyn A.M., Syerov Yu. // *Proceedings of the International Conference on Computer Science and Information Technologies, CSIT`2006, September 28th-30th, Lviv, Ukraine.* – P.56–58. 2. Сєров Ю.О. Моделювання поведінки та класифікація учасників Веб-спільнот на основі нечітких множин // *Вісник Національного університету “Львівська політехніка”*. – 2008. – № 610 – С. 218–228. 3. Сєров Ю.О. Використання нечітких множин для моделювання активності учасників веб-спільнот // Р.Б. Кравець, Ю.О. Сєров / *Вісник Харківського національного університету радіоелектроніки “АСУ та прилади автоматики”*, вип. 141, 2007. – С. 113–118. 4. Круглов В.В. Нечеткая логика и искусственные нейронные сети: Учеб. пособие / Круглов В.В., Дли М.И., Голунов Р.Ю. // – М.: Издательство физико-математической литературы, 2001. – 224 с. – ISBN 5-94052-027-8. 5. Welser H. T. Visualizing the Signatures of Social Roles in Online Discussion Groups / Howard T. Welser, Eric Gleave, Danyel Fisher, Marc Smith // <http://www.cmu.edu/joss/content/articles/volume8/Welser/>. 6. Brandes U. Exploratory Network Visualization: Simultaneous Display of Actor Status and Connections / Ulrik Brandes, Jörg Raab, Dorothea Wagner // [web site] <http://www.cmu.edu/joss/content/articles/volume2/BrandesRaabWagner.html>. 7. Freeman L. Visualizing Social Networks, / Linton C. Freeman, [web site] <http://www.cmu.edu/joss/content/articles/volume1/Freeman.html>. 8. Сєров Ю.О. Аналіз комунікативних процесів у Веб-спільнотах середовища Веб 2.0 / Сєров Ю.О., Пелецишин А.М., Слобода К.О. // *Східно-Європейський журнал передових технологій* № 1/2 (37) / 2009. – Харків. – 2009. – С.38–41. 9. Pareto 80/20 Principle Rule <http://www.accountingissue.info/pareto-80-20-principle-rule.html>. 10. Пелецишин А.М. Позиціонування сайтів у глобальному інформаційному середовищі (монографія) / А.М. Пелецишин. – Львів: Вид-во Нац. ун-ту “Львівська політехніка”, 2007. – 260 с. 11. Page, The PageRank citation ranking: Bringing order to the Web / Page, Lawrence; Brin, Sergey; Motwani, Rajeev and Winograd, Terry // [web site] <http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf>.