

**Дослідження методів пошуку маркетингових даних в мережі Internet****М'якшило О. М., Джуренко Т.С.***Національний університет харчових технологій*

На розвиток маркетингу, як і на інші економічні дисципліни, мають вплив нові технології. Якщо на початку ХХ століття маркетинг розглядався як дистрибуція, потім, як управління продажами і на кінець, як управління торговими марками, то на сьогодні, розвиток інформаційних технологій міняє характер маркетингової діяльності компанії. [1]

Дослідницька (аналітична) функція маркетингу являє собою фундамент усієї маркетингової діяльності підприємства та охоплює всю підготовку, що передуює прийняттю будь-якого рішення щодо кожного елементу комплексу маркетингу. Без ґрунтовного дослідження стану та перспектив розвитку зовнішнього середовища, без аналізу внутрішнього середовища підприємства практично неможливо провадити господарську та комерційну діяльність, змінювати на користь підприємства контрольовані чинники зовнішнього середовища. Постійний контроль змін середовища вимагає створення налагодженої системи спостереження, що ґрунтується на ефективних методах збору та аналізу ринку. Своєчасне надходження інформації про зміни в маркетинговому середовищі дозволить обмежити підприємство від потрясінь, що стане запорукою її стійкості та конкурентоспроможності.

Мета цього дослідження – дослідити та проаналізувати існуючі пошукові системи з точки зору їх ефективності для збору маркетингової інформації в мережі Інтернет, підвищити ефективність пошукових процесів у глобальній мережі.

Для реалізації потреби в пошуку релевантної маркетингової інформації пропонується, на основі діючої інформаційно-пошукової системи, застосувати метод контент-аналізу та семантичний вебпошук для створення контекстозалежного пошуковика.

Останнім часом в технології пошуку все частіше стали впроваджуватися елементи контент-аналізу, методологія якого виникла наприкінці ХІХ - початку ХХ вв. Розрізняють кількісний і якісний контент-аналіз. Якщо якісний контент-аналіз базується на глибокому лінгвістичному та семантичному аналізі окремих пропозицій і всього тексту, то основою кількісного контент-аналізу є статистичні підходи.

Отримали розвиток такі напрямки контент-аналізу, як "Text Mining" і "Web Mining", які припускають автоматичне виявлення нового сенсу з текстових масивів, нових даних, феноменів, фактів - знань. Все частіше виникають спроби залучення методів контент-аналізу, а точніше Text Mining в реальні пошукові системи. І ці спроби не даремні - вони обумовлені обсягами і темпами зростання Інтернет [2].

Це інструмент, який дає можливість аналізувати великі обсяги інформації у пошуках тенденцій, шаблонів і взаємозв'язків, здатних допомогти у прийнятті

стратегічних рішень.

У Text Mining з'явилися нові можливості: автоматичне реферування текстів та виявлення феноменів, тобто понять і фактів. Можливості сучасних систем Text Mining можуть застосовуватися у системах управління знань для виявлення шаблонів у тексті, для розподілу інформації за профілями, створення оглядів документів. Text Mining забезпечує новий рівень семантичного пошуку документів.

Онтологічний підхід – це створення методології побудови онтологій та їх використання при аналізі, відібраних з Інтернету, текстів документів. Характерною рисою цього підходу є поділ предметної області на класи об'єктів і визначення їх фундаментальних властивостей та правил, що визначають їхні зміни і поведінку.

Латентно-семантичний аналіз (ЛСА) доцільно використовувати для обробки текстів - результатів пошуку, в якості практичного методу, що характеризує значення ключових слів. Ці кореляції моделюють механізм мислення людини, зіставляючи частини тексту за змістом. ЛСА використовує матрицю сингулярного розкладання на основі ермітової матриці, яка описує використання слів в документах. Нехай стовпці матриці відповідають документам, а рядки - словам, що зустрічаються в документах. Елементи матриці тоді являють собою кількість вживань даного слова в даному уривку:

$$t_i^T \rightarrow \begin{matrix} & d_j \\ & \downarrow \\ \begin{bmatrix} x_{1,1} & \dots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \dots & x_{m,n} \end{bmatrix} \end{matrix}$$

Якщо з усіх сингулярних значень відібрати  $k$  найбільших, то ми одержимо апроксимацію вихідної матриці матрицею рангу  $k$ . Важливо зазначити, що результати методу ЛСА, залежать не тільки від частотності використання слів в уривках. Метод ґрунтується на виявленні більш глибоких («латентних») зв'язків і, таким чином, краще моделює людське сприйняття тексту ніж прості методи, засновані на частотності вживання слів [3].

Отже, розроблення і використання комбінованої інформаційної технології контекстного пошуку може забезпечити отримання маркетингової інформації з мережі Інтернет більшої релевантності, без додаткових витрат за менший проміжок часу; віднайти латентні (приховані) дані, які традиційним способом зробити дуже важко і ресурсовитратно.

#### Література

1. *Котлер Ф. І.* Маркетинг менеджмент. : навч. посіб. / Ф.І. Котлер, Л. А. Волковой, Ю. Н. Каптуревський. – СПб.: Питер, 2000. – 752 с..
2. *Шалак В. И.* Контент-анализ: навч. посіб. / В. И. Шалак – Приложения в области: политологии, психологии, культурологии, экономики, рекламы. — Российская акад. наук; Ин-т философии. — М. : Омега — Л, 2004.
3. *Башмаков А. И.* Интеллектуальные информационные технологии: Учеб. пособие. / А.И.Башмаков, И. А. Башмаков – М.: МГТУ им. Н.Э.Баумана, 2005. – 304 с.