

**К.О. ТАРАСЕНКО, С.В. ОРЕХОВ**, канд. техн. наук, доцент

### **Розробка програмного забезпечення для дослідження множини ключових слів на прикладі ринку побутової техніки та електроніки**

Інформація, яку залишають користувачі в мережі Інтернет несе в собі маркетингову цінність. А саме на основі неї можна виявити нові події чи факти, які так чи інакше впливають на ринкову ситуацію.

Останні дослідження підтвердили той факт, що інформаційні потоки відображають ринкові події [1]. Тому виникає необхідність аналізувати інформаційні потоки зокрема потоки Інтернет новин, повідомлення користувачів у соціальних мережах, відгуки у дошках об'яв про той чи інший товар або послугу і так далі. В роботі пропонується реалізувати перший етап такого аналізу – створення словника ключових слів, що описують ринкову подію на основі відгуків у дошці об'яв. У якості прикладу розглянуто ринок побутової техніки та електроніки.

Словник, що створюється, включає множину ключових слів, яка динамічно змінюється в залежності від ринкових подій у часі. У якості критерію відбору ключових слів, що описують ринкові події була обрана міра TF-IDF [2].

Після розрахунку даного показника наступним кроком є лематизація слів, тобто приведення їх до нормальної (словникової) форми.

На основі отриманих результатів обирається множина ключових слів, які мають найбільше значення TF-IDF.

Далі здійснюється пошук зв'язку між словами. Для цього використовуються асоціативні правила (market basket analysis) [3] та сервіс Wordstat пошукової системи Яндекс.

Результатом роботи є розроблене ВЕБ орієнтоване програмне забезпечення у вигляді дошки об'яв з використанням компоненти «хмара тегів». Цей компонент візуалізує перші десять комбінацій ключових слів. Кожна комбінація описує множину ключових слів, які відображають попит на ринку побутової техніки та електроніки Харкова.

#### **Список літератури:**

1. Черенков І.А. Обоснование прогнозирования цен полимеров посредством новостного потока / И.А. Черенков, С.В. Орехов // Східно-Європейський журнал передових технологій. – Харків : Технологічний центр, 2010. – №5/7 (47). – С. 18-21.
2. What does tf-idf mean? // <http://www.tfidf.com/>, 13.11.13.
3. Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Технологии анализа данных. Data Mining, Visual Mining, Text Mining, OLAP – 2 изд., перераб. и доп. – СПб.: БХВ - Петербург, 2007. – С.124-142.