

ІДЕНТИФІКАЦІЯ КОРИСТУВАЧІВ В РІЗНИХ СОЦІАЛЬНИХ МЕРЕЖАХ

Активне зростання аудиторії соціальних медіа в мережі Інтернет, таких, як соціальні мережі, форуми, блоги та інтернет-ЗМІ, привів до становлення цих ресурсів як нового джерела даних для прийняття маркетингових рішень. Аналіз соціальних даних стрімко набирає популярність у всьому світі [1–3] завдяки появі в 1990-х роках онлайн-сервісів соціальних мереж (LiveJournal, Facebook, Twitter, YouTube та ін.). З цим пов'язаний феномен соціалізації персональних даних: стали публічно доступними факти біографії, листування, щоденники, фото, відео, аудіо-матеріали, нотатки про подорожі і т.п. Таким чином, соціальні мережі є унікальним джерелом даних про особисте життя і інтереси реальних людей. Це відкриває безпрецедентні можливості для вирішення дослідницьких і бізнес-задач (багато з яких до цього неможливо було вирішувати ефективно через брак даних), а також створення допоміжних сервісів і додатків для користувачів соціальних мереж.

Однією з фундаментальних проблем при використанні соціальної інформації про користувача є її фрагментарність. Серед множини різних онлайн-соціальних мереж щороку з'являється безліч як універсальних, так і нішевих соціальних сервісів, і для активних користувачів Інтернет типово мати кілька профілів в різних соціальних мережах. Незважаючи на те, що існують спроби щодо забезпечення єдиного способу взаємодії між різними соціальними платформами (наприклад, OpenSocial), вони не набули широкого застосування, а нові соціальні сервіси продовжують з'являтися. Ідентифікація користувача в різних соціальних мережах дає змогу отримати більш повну картину про соціальну поведінку даного користувача в мережі Інтернет. Виявлення акаунтів, що належать одній людині, в декількох соціальних мережах, дозволяє отримати більш повний соціальний граф, що може бути корисним в багатьох задачах, таких, як інформаційний пошук, інтернет-реклама, рекомендаційні системи і т.п.

Оскільки пошук акаунтів користувача в різних мережах в загальному випадку вимагає наявності актуальних даних про всіх користувачів цих мереж, доцільно обмежити простір пошуку найближчими сусідами будь-якого користувача, акаунти якого в досліджуваних мережах відомі.

Таким чином, завдання ідентифікації користувачів в різних соціальних мережах в локальній перспективі передбачає зіставлення

акаунтів користувачів в рамках списків контактів деякого центрального користувача в різних соціальних мережах. Таке завдання часто виникає при роботі з контактами користувачів в соціальних мета-сервісах, які, зокрема, можуть служити для об'єднання потоків новин в соціальних сервісах або надання єдиної системи обміну повідомленнями. Подібна задача виникає також при використанні функції автоматичного об'єднання контактів з різних джерел (телефонна книга, соціальні мережі, месенджери), поширеної в сучасних мобільних пристроях.

Необхідним інструментарієм для дослідження виступають додаткові програмні продукти, що дозволяють маніпулювати потоками даних і організувати їх максимально зручним для подальшого аналізу способом.

Програмне забезпечення для аналізу даних соціальних мереж [4, 5] зазвичай використовує теорію графів для дослідження соціальних структур, як аналітично, так і візуально. Основні конструкції – це вузли (сутності, у яких ми зацікавлені, як правило, люди), а також ребра, які з'єднують їх. Багато продуктів існує з відкритим вихідним кодом, які можна використовувати безкоштовно в некомерційною умовою.

Після проведення аналізу існуючого програмного забезпечення можна зробити висновок:

- більшість програмних продуктів дозволяє лише побудувати граф даних з соціальної мережі;
- на ринку програмного забезпечення відсутні продукти, які дозволяють проводити аналіз профілів користувачів соціальних мереж з метою інтернет-маркетингу.

Таким чином, пропонується розробка власного програмного забезпечення, яке дозволяє ідентифікувати користувача соціальних мереж.

Список літератури

1. Najork M., Wiener J.L. *Breadth-first crawling yields high-quality pages* // *Proceedings of the 10th international conference on World Wide Web.* – ACM, 2011. - P. 114-118.
2. *Number of social media users worldwide from 2010 to 2020* // <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>, 11.09.2016.
3. *World map of social networks research by Alexa and SimilarWeb* // <http://vincos.it/world-map-of-social-networks/>, 14.09.2016.
4. *Semantic Text Analysis APIs Using Natural Language Processing* // <http://www.alchemyapi.com/>, 15.11.2016.