УДК 004.056

Бісюк В.А., Новожилова О.А.
Центральноукраїнський національний технічний університет

Методи ідентифікації спаму у соціальних мережах

Сучасні соціальні мережі, набувають все більшої популярності серед користувачів Інтернету, їх різновиди та функціонал активно збільшуються. Вони об'єднують в собі блоги, сховища медіа-контенту, записи персональної інформації (Facebook, Вконтакте), wiki-подібні енциклопедії та інші. Фактично Web-сайти соцмереж являють собою велике сховище загальнодоступної інформації, в першу чергу, персонального характеру і надають можливість комунікації та обміну даними, як окремих користувачів, так і груп або співтовариств [1], об'єднаних спільними інтересами, вподобаннями тощо.

Такий швидкий розвиток та популярність соцмереж призвело до поширення одної з основних проблем користувачів – надлишку небажаної інформації «спаму».

Спам є масштабною розсилкою комерційної, політичної та іншої реклами (інформації) або іншого виду повідомлень особам, які не виказували бажання їх отримувати. Більшість способів спам-атак засновані на методах соціальної інженерії (залучення користувачів недобросовісною або навіть шкідливою рекламою і т.д.), а також на використанні вразливостей в механізмах роботи соціальних мереж. Технології розсилки спаму в соціальних мережах удосконалюються: спамери визначають користувачів соцмережі на фотографіях та відеозаписах, додаються в друзі, запрошують в групи і так далі, в цілому використовують всі можливості соцмережі в корисливих цілях.

Протидія спамерам в соціальних мережах важлива для поліпшення сервісів, що надаються соціальною мережею для учасників, зменшення кількості небажаного і небезпечного контенту.

Переважно в соцмережах використовують декілька різновидів анти-спам стратегій:

- засновані на ідентифікації (Identification-based),
- засновані на ранжируванні (Rank-based),
- засновані на інтерфейсі і обмеженнях (interface-based, limit-based).

Але найкращі результати досягаються при поєднанні всіх стратегій в єдиному комплексі [2].

Головною причиною масових розсилок спаму є наявність спамерів і різних спамерських пошукових роботів в соцмережах, тому їх швидка і якісна ідентифікація є запорукою успіху систем захисту, що базується на аналізі профілів користувачів, соціальної поведінки, повідомлень, що розсилаються користувачами та контенту, який додається і створюється ними власноруч. Аналіз цих параметрів дозволяє визначити особливості, які притаманні поведінці звичайного користувача, а отже відповідно може поліпшити

розпізнавання шкідливих спам-ботів і спамерів. Якісна система ідентифікації повинна враховувати всі ознаки, які потенційно можуть бути важливі для класифікації користувачів.

Для проведення класифікації необхідна перевірка обробка профілів користувачів, яка в першу чергу включає в себе безпосереднє виділення інформації, ознак і атрибутів з Web-сторінки облікового запису [3].

Модель «Множина слів» (Bag of words) - використовується в обробці природної мови і пошуку інформації (information retrieval). У цій моделі, текст (речення або документ) представляється як невпорядкований набір слів, без врахування граматики і порядку слів. Для цього необхідно прибрати HTML-теги, знаки пунктуації, стоп-слова, все слова перевести в нижній регістр. Стоп-словами є слова, які не несуть самостійного смислового навантаження. Як правило, до них відносять прийменники, сполучники, частки, займенники, вступне слово, вигуки і т.д.

Далі, для виділеної з профілю користувача інформації, необхідний стемінг - процес знаходження основи слова для заданого вихідного слова.

Одним з найбільш популярних і ефективних алгоритмів стемінгу є Стеммер Портера. Його головною перевагою є те, що не використовуються словники, і виділення основи здійснюється шляхом перетворення слова згідно певних правил. Алгоритм не використовує баз основ слів, а лише, застосовує послідовно ряд правил, відсікає закінчення і суфікси, ґрунтуючись на особливостях мови, в зв'язку з чим працює швидко, але не завжди безпомилково [4].

Подальша обробка включає в себе виділення атрибутів (довжина повідомлень, їх повторюваність та кількість, характер медіа-контенту і т.д.) за якими буде визначатись «благонадійність» профілю користувача та степінь їх релевантності. Потім на основі аналізу цих атрибутів системою ідентифікації приймається рішення про класифікацію профілю. Подальші дії залежать від вимог адміністрації ресурсу – попередження, тимчасове блокування або видалення профілю.

Впровадження та вдосконалення таких систем дозволить значно зменшити навантаження на канали зв’язку та сервери соцмереж, підвищити загальну комфортність користування соцмережами і захистить користувачів від небажаної або шкідливої інформації.

Список використаних джерел

1. *Boyd, D. Social Network Sites: Definition, History, and Scholarship / D. Boyd, N. Ellison // Journal of Computer-Mediated Communication. – 2007. – Vol. 13, № 1. – P. 210–230.*
2. *Webb, S. Uncovering Social Spammers: Social Honeypots+Machine Learning / S. Webb , J. Caverlee, K. Lee // Proceedings of the 33rd Annual ACM SIGIR Conference (SIGIR 2010), 19–23 July 2010, Geneva, Switzerland. – Geneva, 2010.*
3. *А.А. Куликова. Подход к классификации пользователей в социальных сетях. / А.А. Куликова // Информационные технологии. - 2011 - 3/2 (51) – С. 14-18.*
4. *Електронний ресурс / https://ru.wikipedia.org/wiki/Стеммер_Портера.*