

## **ФІЛЬТРАЦІЯ СПАМУ ІЗ ЗАСТОСУВАННЯМ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ**

**Никитина Л.А., Радіонов О.С.**

*Національний технічний університет  
«Харківський політехнічний інститут»,  
м. Харків*

Спам – масова розсилка рекламних повідомлень, що проводиться без згоди одержувача. Зазвичай вона здійснюється не централізовано, з одного сервера, а за допомогою так званої спамерської мережі, одночасно з багатьох комп'ютерів, підключених до інтернету. Негативними наслідками спаму є: 1) у спамі часто містяться віруси (зазвичай боти, які включають ваш комп'ютер в існуючу спамерську мережу, що, в свою чергу, може привести до його потрапляння в чорний список); 2) спам-розсилки генерують величезний трафік, створюючи серйозне навантаження на мережеві ресурси; 3) в спам-розсилках часто зустрічаються шахрайські листи (наприклад, "нігерійські листи" та фішинг).

Сучасні методи боротьби зі спамом значно відрізняються від примітивних поштових фільтрів. У них використовуються технології машинного навчання і методи глибокого аналізу мільярдів листів, що дозволяють вивчати роботу поштових серверів як єдиний процес.

У проєкті, що розробляється, запропоновано виконувати двоетапну фільтрацію спаму: на першому етапі отримані повідомлення попередньо класифікуються; на другому етапі застосовується спам-фільтр, налаштований на кінцевого користувача. Для попередньої класифікації було розроблено модель спам-повідомлення. Для класифікації повідомлень виконуються такі кроки: у тексті повідомлення виділяються лексеми, з множини яких формуються атрибути класів повідомлень; на основі атрибутів виконується побудова семантичної мережі повідомлення; за результатами аналізу семантичної мережі формується рішення про приналежність повідомлення до певної категорії. При побудові семантичної мережі пропонується не тільки оцінювати частоту появи лексем в повідомленні, але і визначати міру близькості лексем в рамках прийнятої мінімальної семантичної структури повідомлення (речення). Для виконання класифікації потоків розроблена штучна нейронна мережа із застосуванням алгоритму зворотного поширення помилки. На другому етапі використовується спам-фільтр кінцевого користувача, з налаштованим профілем, який являє собою набір уточнених векторів атрибутів спам-повідомлень та легітимних повідомлень.

Запропонований спосіб класифікації повідомлень дозволяє враховувати переваги кінцевих користувачів при фільтрації спаму на поштовому сервері.