Загрози інформаційній безпеці у рекомендаційних системах соціальних медіа

УДК 004.9

Єлизавета Мелешко

Центральноукраїнський національний технічний університет,
elismeleshko@gmail.com

Основною сферою використання рекомендаційних систем на сьогоднішній день є маркетинг товарів та послуг, а також інформаційний пошук.

Рекомендаційні системи (РС) – програмне забезпечення, що використовується для прогнозування вподобань користувача веб-сайту на основі зібраної про нього раніше інформації.

РС в електронній комерції виконують завдання по збільшенню відсотка продажів товарів та послуг, а також скороченню часу пошуку потрібних товарів та послуг відвідувачем сайту. РС на сайтах контент-проектів покликані збільшувати час перебування відвідувача на сайті, глибину перегляду і залученість. РС в пошукових роботах застосовуються для підвищення релевантності пошукової видачі для користувача.

Метою даної роботи є дослідження загроз інформаційній безпеці у рекомендаційних системах соціальних медіа та визначення ефективних методів протидії ним.

За методами збору даних рекомендаційні системи можна класифікувати на:

1. РС з явним збором даних. Користувач добровільно надає системі необхідні для її роботи дані. Явний збір даних може полягати в проханні користувачу поставити диференційовану оцінку тому чи іншому об’єкту, створити список «улюблених» об’єктів або заповнити анкету, пройти тест. Основний недолік – користувачу необхідно здійснювати деякий набір дій.

2. РС з неявним збором даних. Збір даних відбувається за допомогою спостереження за поведінкою користувача – оцінками, покупками, переходами за посиланнями тощо. Перевагою методу є те, що для збору даних користувачу не треба здійснювати ніяких додаткових дій, все що йому потрібно – просто користуватися веб-сайтом в своєму звичайному режимі. Але даний метод підіймає ряд етичних питань, напр., таких як приватність даних.

3. РС, що поєднують обидва типи збору даних.

Серед властивостей рекомендаційних систем, що є важливими для інформаційної безпеки користувачів можна виділити наступні:

1. Приватність користувача. РС збирають велику кількість даних про користувачів, значну частину яких користувачі охоче надають самі в обмін на корисні рекомендації. Однак для більшості користувачів, важливо щоб їхні вподобання залишалися приватними, тобто, жодна третя сторона не могла використовувати РС, щоб дізнатися інформацію про них або їх вподобання. Дана загроза цілком реальна. Як один з прикладів можна навести скандальні ситуації с рекомендаціями друзів у Facebook, які виникали через експерименти з використанням даних геолокації, подібні рекомендації частково порушували приватні дані людей та давали інформацію третім особам про їх переміщення.

Для забезпечення приватності користувача РС можна застосовувати

наступні методи:

– Інформування користувачів про те, яку інформацію про них збирає РС, гнучкі налаштування параметрів конфіденційності.

– Анонімізація – інформація про користувача може частково видалятися або піддаватися обфускації (маскуванню) користувачем або власником РС.

– Рандомізація – дані користувача (напр., виставлені об'єктам оцінки) можуть бути частково зашумлені випадковими значеннями. Необхідний рівень шуму залежить від того, як часто дані будуть використовуватися, і передбачає балансування між точністю прогнозування та конфіденційністю користувача.

– Шифрування даних користувача, що зберігаються в базі даних РС.

2. Ризик для користувача. В деяких випадках рекомендації можуть бути пов'язані з ризиком. Напр., якщо об'єктами в РС є акції, кредити, депозити, ліки, медичні послуги, політичні акції тощо. В таких випадках може бути необхідним врахування не тільки вподобань користувача при формуванні рекомендацій, а й інших факторів, врахування яких здатне мінімізувати ризик для користувача, що буде переглядати та обирати рекомендації.

3. Робастість системи до атак. Здатність системи надавати адекватні рекомендації при появі некоректної інформації. Некоректна інформація може виникати при атаках на РС з метою збільшення рейтингу на певні об'єкти, напр., при створенні великої кількості фейкових акаунтів та ботів, які виставляють високі оцінки певному об'єкту чи об'єтам.

Побудова робастних РС базується на двох принципах: 1) виявлення спаму серед дій користувачів; 2) невраховування при побудові рекомендацій даних користувачів, що поширюють спам. На сьогоднішній день існують методи, які дозволяють виявити атаку на РС, використовуючи той факт, що розподіл коефіцієнтів подоби користувачів змінюється при наявності спам-користувачів в РС. Так як при атаці створюють не один фейковий профіль, а декілька схожих, такі спам-користувачі будуть мати незвично високу схожість, у порівнянні зі звичайними користувачами. Однак, надійні методи виявлення атак на РС та протидії ним все ще залишаються активною областю досліджень.

4. Бульбашка фільтрів. Класичні РС пропонують користувачам об'єкти, виходячи лише з їх попередніх вподобань. Отже, користувач потрапляє в інформаційне середовище, в якому спостерігає лише обмежену кількість однотипних об'єктів. Бульбашка фільтрів викликає наступні загрози:

– Користувач не одержує альтернативну інформацію, яка може бути йому корисною та ефективніше вирішить задачі його пошуку.

– У користувача формується викривлена точка зору на інформаційне середовище, так як він не бачить картини в цілому.

Отже, внаслідок бульбашки фільтрів користувачі можуть бути частково дезінформовані, що може призводити до негативних наслідків, якщо, напр., РС застосовується на суспільно-політичному новинному сайті.

Для вирішення проблеми бульбашки фільтрів, як правило, застосовують висування додаткових вимог до РС, напр., забезпечення властивостей РС: різноманітність (Diversity), неочікуваність (Serendipity), новизна (Novelty). При внесенні відповідних змін в роботу РС, буде зменшуватися точність прогнозування, але можна буде подолати проблему бульбашки фільтрів.