Fake News та захист інформації

УДК 004.046.8249

Томайли Дмитро

аспірант, Національний авіаційний університет, dmytro.tomayly@gmail.com

В наш час існує велика кількість різноманітних наукових праць на тему захисту інформації. Частина цих робіт суто теоретична, інша частина є основою комерційних продуктів захисту інформації. Але, всі вони фокусуються на кінцевій проблемі, такій як: захист корпоративного веб додатку, листування, мобільного пристрою, інтелектуальної власності та інше. Не існує всебічного та універсального засобу захисту інформації, який б зміг надати захист від будь-яких векторів атак.

Провівши всебічний аналіз проблематики захисту інформації в корпоративному сегменті ринку, було проведено декілька вузьких досліджень кінцевих проблем (результати досліджень можна знайти у збірнику матеріалів IV міжнародної науково-практичної конференції «Сучасні тенденції розвитку науки» від грудня 2018 року (ISBN 978-617-7640-39-3)). Під час досліджень було виявлено, що не існує проблеми реалізації додатку, який дозволить зловмиснику отримати несанкціонований доступ до будь-якої інформації. Також, під час дослідження, в якості перевірки концепції (POC), були розроблені тестові додатки, які дозволяють отримати необхідну інформацію. І отримати цю інформацію стало можливим, навіть якщо були застосовані всі можливі засоби захисту від несанкціонованого доступу до інформації. Таким чином вдалося, не тільки перевірити можливість отримати несанкціонований доступ до інформації, але й довести, що реалізація подібних шкідливих програмних засобів не потребує великих затрат часу або ресурсів.

Серед висновків проведеного аналізу та досліджень можна виділити наступне: 1) складність сучасних програмних засобів дуже велика, тому завжди можна знайти вразливість, яку можна використати для несанкціонованого доступу до інформації; 2) розвиток сучасних додатків призводить до використання все більшої кількості вже готових алгоритмів\бібліотек\фреймворків, що унеможливлює перевірку всього вихідного коду кінцевого додатку; 3) виходячи зі складності та, практично, неможливості надати всебічний захист, необхідно частково змінити концепцію підходу до захисту інформації.

Якщо проаналізувати найбільші фінансові втрати корпорацій, спричинені витоками інформації, за останній час, то можна прийти до висновку, що найбільші фінансові та репутаційні втрати спричинені не самим витоком інформації або використанням цієї інформації, а безпосередньо самою новиною про факт витоку або факту нецільового використання додатку\сервісу. Серед найгучніших прикладів можна навести: 1) перетворення Big 5 у Big 4, де одна з 5-х великих аудиторських компаній практично припинила своє існування завдяки чуткам, що відбувся виток інформації про банкрутство компанії, спричинене розслідуванням комісії по цінним паперам та біржам; 2) падіння ціни акцій компанії Facebook, завдяки чуткам про виток інформації, під час розслідування про причетність компанії до перешкоджання або маніпуляцію волевиявлення громадян США під час виборів президента країни; 3) падіння

ціни акцій компанії Intel, завдяки витоку інформації про приховування інформації щодо знайдених в 2018 році вразливості «Meltdown» та «Spectre». Як можна побачити, сам виток інформації, навіть в тих випадках коли він мав місце, не спричинив істотних втрат, на відміну від недостовірної інформації про компанію або обставини розслідувань, призвела до великих фінансових та репутаційних втрат. При чому, можна переконатися, що наслідком репутаційних втрат є фінансові втрати, не тільки на прикладі корпорацій. Цю залежність легко знайти і по відношенню до цілих країн, таким прикладом може бути Російська Федерація, коли чутки про санкції призводять до більш суттєвих фінансові втрат, ніж самі санкції.

Важливо підкреслити, що в наш час недостовірна інформація більш шкідлива, чим несанкціонований доступ до інформації. Таким чином, необхідно продовжувати дослідження, безпосередньо, у напрямку виявлення недостовірної інформації завдяки автоматизованим технічним засобам.

Провівши аналіз існуючих алгоритмів контекстного аналізу текстової інформації, можна сказати, що найбільш цікавою, для подальшого дослідження, є використання згорткових нейронних мереж (ЗНМ). Під час аналізу також було проведене порівняння точності класифікації реакції людей на новину у вигляді відгуків користувачів відомої соціальної мережі (позитивна або негативна реакція). Порівняння ЗНМ з так званим «нейронним мішком слів» (NBoW), який найчастіше використовується для подібного аналізу, можна зайти у таблиці 1.

Таблиця 1

Порівняння точності розпізнавання реакції користувачів

Алгоритм	Точність класифікації
NBoW	65.3%
ЗНМ, ручний підбір ознак	77.5%
ЗНМ, кращій результат	83%

Як можна побачити, ЗНМ має досить високу точність на поточній виборці та при очікуваному простому результаті (1 – позитивна реакція, 0 – нейтральна реакція, –1 – негативна реакція). При зростанні складності текстів у тестовій виборці для обох алгоритмів значно падає точність класифікації.

Висновком з проведеного дослідження є необхідність вдосконалення внутрішніх алгоритмів реалізації роботи ЗНМ, які потребують додаткового дослідження. Але вже зараз можна виділити основні моменти: 1) ЗНМ значно виграє у складності реалізації, порівнюючи з іншими алгоритмами нечіткої логіки; 2) Необхідно аналізувати текст цілими реченнями або групами речень тому, що в сучасних мовах багато речових оборотів, які не повинні розглядатися буквально, наприклад, в залежності від контексту фразеологізми, на кшталт «piece of cake», «Проще пареной репы» та «простіше простого» система буде сприймати невірно. На поточний момент дослідження можна стверджувати, що необхідний ступінь точності, в загальному випадку, можливо буде досягнути лише шляхом комбінації різноманітних алгоритмів нечіткої логіки. Наприклад, комбінації експертної системи та ЗНМ, де експертна система дозволить первинно надати інформацію про тлумачення речових оборотів, а нейронна мережа, в свою чергу, дозволить надати необхідну класифікацію.