УДК 004.032.26:004.056]:004.773.3

ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИЯВЛЕННЯ ТА КЛАСИФІКАЦІЇ ШАХРАЙСЬКИХ ПОВІДОМЛЕНЬ У ЕЛЕКТРОННІЙ ПОШТІ

Вельма І.Ю.
Науковий керівник – д. т. н., проф. Мартинчук О. О.
Харківський національний університет радіоелектроніки, каф. ІКІ
м. Харків, Україна
e-mail: ihor.velma@nure.ua

In today's digital world, email phishing has become a serious threat to users and organizations. Malicious actors employ various methods and techniques to deliver fraudulent messages, attempting to deceive recipients and gain access to their confidential data or financial resources. Consequently, there is a need for effective methods of detecting and classifying phishing emails.

Використання нейронних мереж є одним з перспективних напрямків у розв'язанні проблеми виявлення та класифікації шахрайських повідомлень у електронній пошті. Нейронні мережі, зокрема глибокі нейронні мережі, є потужним інструментом у сфері кібербезпеки завдяки їх здатності аналізувати великі обсяги даних та виявляти складні зв'язки та патерни. Глибокі нейронні мережі складаються з багатьох шарів нейронів, які обробляють вхідні дані на різних рівнях абстракції. Це дозволяє їм автоматично виявляти навіть тонкі та складні закономірності в даних, які можуть бути важко виявити за допомогою традиційних методів. Наприклад, глибокі нейронні мережі можуть аналізувати текстові повідомлення у електронних листах та виявляти підозрілі ключові слова, фрази або шаблони, що вказують на можливість шахрайства. Однією з переваг використання нейронних мереж є їх здатність до самонавчання. Це означає, що мережа може вдосконалювати свої навички на основі нових даних, які вона отримує, що дозволяє підтримувати високий рівень точності та ефективності виявлення шахрайських повідомлень навіть у змінному середовищі. Таким чином, використання нейронних мереж, зокрема глибоких нейронних мереж, є потужним інструментом для виявлення та класифікації шахрайських повідомлень у електронній пошті, і цей підхід може значно полегшити роботу при боротьбі з кіберзагрозами в цій сфері.

Для виявлення та класифікації шахрайських повідомлень застосовуються різноманітні архітектури нейронних мереж, такі як згорткові нейронні мережі (CNN), рекурентні нейронні мережі (RNN) та їхні комбінації. Методи передбачення на основі нейронних мереж можуть використовувати як текстові дані повідомлень, так і додаткові характеристики, такі як адреси електронної пошти, метадані повідомлень та інші атрибути.

Основна ідея згорткових нейронних мереж (CNN) полягає у тому, щоб вони могли автоматично вивчати корисні функції (ознаки) з вхідних даних. Це відбувається за рахунок використання фільтрів (які також називають ядром або маскою), які згортаються по всій вхідній матриці даних для виділення особливостей. У процесі навчання мережі ваги цих фільтрів автоматично оптимізуються таким чином, щоб вони могли ефективно розпізнавати корисні ознаки. Однією з ключових переваг згорткових нейронних мереж є їхній потенціал для автоматичного виявлення локальних шаблонів у вхідних даних [1].

Основна ідея рекурентних нейронних мереж полягає у використанні зворотного зв'язку, що дозволяє інформації передаватися з одного кроку до наступного. Кожен часовий крок рекурентної мережі обчислює вихід на основі введених даних та попереднього стану. Це дозволяє їм враховувати контекст та історію даних при роботі з послідовностями. Однією з ключових переваг рекурентних нейронних мереж є їх здатність до моделювання довгострокових залежностей у послідовностях даних [1].

Дослідження використання нейронних мереж для виявлення та класифікації шахрайських повідомлень показали високу ефективність цих методів. Нейронні мережі здатні автоматично виявляти характерні ознаки шахрайських повідомлень та відокремлювати їх від легітимних. Застосування нейронних мереж дозволяє знизити кількість фальшивих позитивів та підвищити точність виявлення шахрайських повідомлень.

Додатковою перевагою використання нейронних мереж є їхня здатність до адаптації до нових видів шахрайства та змін у технологіях шахрайства. Це робить їх ефективним інструментом для захисту від постійно змінюючихся загроз у сфері електронної пошти.

Використання нейронних мереж для виявлення та класифікації шахрайських повідомлень у електронній пошті виявляється як обіцяний підхід, що дозволяє підвищити ефективність захисту користувачів та організацій від кіберзлочинності. Результати досліджень свідчать про високу ефективність цих методів та їхню здатність адаптуватися до нових загроз, що робить їх важливим інструментом у боротьбі з електронним шахрайством.

Список використаних джерел:

1. McMillan C. The Connectionist Scientist Game: Rule Extraction and Refinement in a Neural Network / C. McMillan, M.C. Mozer, P. Smolensky // Proc. XIII Annual Conf of the Cognitive Science Society, Hillsdale, NJ,
USA. – 2001.