АНАЛІЗ ЗАГАЛЬНОДОСТУПНИХ НАБОРІВ ДАНИХ ДЛЯ ІМІТАЦІЇ МЕРЕЖЕВИХ АТАК

Заковоротний О.Ю., Хулап А.В.
Національний технічний університет "Харківський політехнічний інститут", м. Харків

У сфері виявлення мережевих вторгнень (NIDS) відкриті набори даних відіграють ключову роль, оскільки вони застосовуються для навчання та тестування різноманітних моделей і алгоритмів. Якісні та різнопланові дані дають змогу ефективно перевіряти новітні системи в умовах, наближених до реальних. У цьому контексті важливою задачею є аналіз та порівняння найпоширеніших публічних наборів даних для NIDS, аби визначити, які з них найбільше підходять для подальших досліджень.

У працях [1..3] розглядаються найбільш вживані набори даних, окреслюються їхні особливості, слабкі сторони, а також оцінюється їх придатність для тестування нових систем виявлення вторгнень.

Під час оцінювання ефективності сучасних рішень на основі машинного навчання слід враховувати не лише актуальність даних, але й специфіку мережевого трафіку. Зокрема, використання реальних даних, як у випадку з CICIDS'17 чи LITNET-2020, забезпечує кращу відповідність сучасним загрозам і є цінним для практичного тестування нових рішень.

З огляду на це, доцільно зосередити увагу на наборах даних CICIDS'17, CSE-CIC-IDS2018 та NSL-KDD 2009, які дають змогу ефективно порівнювати методи та тестувати нові підходи на актуальних прикладах. Також варто згадати KDD Cup 1999 — один із найвідоміших наборів для NIDS, створений на основі даних DARPA 1998 року, що містить 41 характеристику та понад 4.8 мільйона записів.

Загальнодоступні набори даних є важливим інструментом для порівняння нових методів із вже відомими та перевіреними підходами до виявлення вторгнень і аномалій. Вони дають змогу об'єктивно оцінити точність, ефективність і продуктивність сучасних рішень, а також сприяють удосконаленню алгоритмів машинного навчання для надійного захисту мереж в умовах обмежених ресурсів.

Література:

1. Mohammadpour, L., Ling, T. C., Liew, C. S., & Aryanfar, A. (2022), "A Survey of CNN-Based Network Intrusion Detection", Applied Sciences, 12(16), 8162, doi: https://doi.org/10.3390/app12168162.

2. Zhen Yang, Xiaodong Liu, Tong Li, Di Wu, Jinjiang Wang, Yunwei Zhao, Han Han. (2022), "A systematic literature review of methods and datasets for anomaly-based network intrusion detection", Computers & Security, Volume 116, doi: https://doi.org/10.1016/j.cose.2022.102675.

3. Chou D., Jiang, M. (2021), "A Survey on Data-driven Network Intrusion Detection", ACM Computing Surveys (CSUR), Volume 54, Issue 9, doi: https://doi.org/10.1145/3472753.