Пулеко І.В., к.т.н., доцент,
Пулеко К.І., студент,
Державний університет «Житомирська політехніка»
Іщенко І.А., ст. викладач,
Свистунович І.В., ст. викладач
Житомирський військовий інститут ім. С. П. Корольова

ОГЛЯД ПРОГРАМ, ЩО ВИКОРИСТОВУЮТЬ МАШИННЕ НАВЧАННЯ ПРИ ВИЯВЛЕННІ АНОМАЛІЙ У КІБЕРБЕЗПЕЦІ

Одним з факторів, що суттєво впливає на час реагування на кібератаки є виявлення і прогнозування нових загроз. Так затримка при реагуванні виникає навіть при загрозах відомих типів. Нові ж види атак, моделі поведінки та інструменти можуть збити фахівців з пантелику, в результаті чого вони будуть реагувати ще повільніше. Гірше того, такі менш помітні загрози, як крадіжка даних, іноді можуть залишитися і зовсім невиявленими. Тому, будуючи захист мереж, доводиться завжди враховувати постійний розвиток технологій, що застосовують зловмисники. На щастя, методи кібератак зазвичай не винаходяться з нуля, а основою для них часто служать тактики, платформи і вихідні коди минулих атак. Саме це і дозволяє застосовувати технології машинного навчання, бо їм є на чому базуватися при накопиченні знань.

Програмне забезпечення (ПЗ) на основі машинного навчання допомагає розпізнати атаку, виявивши спільні риси у новій загрозі та виявлених раніше. Машина, на відміну від людини, проведе таке порівняння швидко - що ще раз підкреслює необхідність застосування адаптивних моделей безпеки. Машинне навчання може полегшити прогнозування нових загроз і скоротити час реагування за рахунок більш ефективної роботи з базою існуючих загроз.

У комп'ютерній безпеці зловмисні проникнення можуть бути виявлені за рахунок незвичайного мережного трафіка або нетипових дій користувача. Такі вторгнення можуть порушити як приватну конфіденційність, так і організаційну. Їхнє виявлення зводиться до аналізу аномалій.

Підходи з використанням методів машинного навчання до проблеми виявлення вторгнень можна поділити на два класи:

1. Виявлення неправильного використання (Misuse Detection) – побудова прогностичної моделі на основі розмічених даних (екземпляри помічені як «нормальні» або «проникнення»). Вони демонструють високу точність виявлення великої кількості відомих атак, але не можуть виявити невідомі та нові атаки.

2. Виявлення аномалій (Anomaly Detection), коли система може виявляти нові атаки як відхилення від «нормальної» поведінки, однак при цьому можливий високий рівень хибних «спрацьовувань», так як виявлені відхилення не завжди являють собою реальне вторгнення.

У доповіді розглядаються саме програмні системи виявлення аномалій (Anomaly Detection Systems, ADS). Основним припущенням ADS є те, що дії зловмисника (події в атакованій системі) обов’язково чимсь відрізняються від поведінки звичайного користувача (від подій в нормальному стані), тобто є аномаліями. Тому такі системи здатні реєструвати і невідомі атаки.

Серед платного ПЗ для рішення задачі виявлення аномалій доцільно було б провести тести та порівняти такі: Numenta, AVORA, Splunk Enterprise, Loom Systems, Elastic X-Pack, Anodot, CrunchMetrics. Однак із-за значної його вартості вдалося провести лише порівняльний огляд по описам та відгукам з Інтернету. Як краще безкоштовне ПЗ для виявлення аномалій розглядалися: Weka Data Mining, Shogun, RapidMiner Starter Edition, Dataiku DSS Community, ELKI, Scikit-learn. Оцінювалися як відгуки користувачів, так і експериментальна якість виявлення аномалій.

Порівняння можливостей та функціоналу ПЗ проводилося за такими показниками:

Якість виявлення аномалій: це найважливіша функція програмного забезпечення виявлення аномалій, оскільки основна мета програмного забезпечення — виявляти аномалії. Програмне забезпечення повинно дозволяти бізнес-користувачам виявляти будь-які незвичайні моделі, поведінку чи події. Тут застосовувалися відомі показники якості виявлення аномалій на основі матриці помилок.

Кількість алгоритмів машинного навчання, що реалізовано в ПЗ. Можливість вибору та налаштування різних алгоритмів машинного навчання дозволяє обирати найкращі для рішення конкретної задачі.

Оповіщення в режимі реального часу: функція виявлення аномалій була б марною без механізму оповіщення користувачів, коли система виявляє аномалію. ПЗ повинно повідомляти користувачів у режимі реального часу, надсилаючи попередження електронною поштою, додатковими програмами та текстовими повідомленнями.

Можливості моніторингу та відображення: програмне забезпечення для виявлення аномалій повинно включати панелі моніторингу та відображення, що налаштовуються, які дозволяють користувачам відображати показники різними способами.

Інтеграція: Найкраще програмне забезпечення для виявлення аномалій повинно легко інтегрується з існуючими системами та поширеними мовами програмування.