Секція 2

ЗБИРАННЯ ТА АНАЛІЗ ІНФОРМАЦІЇ ПРО ВРАЗЛИВОСТІ КОМПОНЕНТІВ І СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ З ВИКОРИСТАННЯМ ЗАСОБІВ BIG DATA

Неретін О.С.
Національний аерокосмічний університет ім. М.Є Жуковського
«Харківський авіаційний інститут»
Науковий керівник Харченко В.С.

Актуальність. Зростання використання систем штучного інтелекту (СШІ) в різних сферах індустрії транспорту, оборони, медицини супроводжується із збільшенням загроз, зростанням кількості і видів кібератак [1]. Враховуючи поширення цих систем в критичних доменах, важливим є забезпечення їх захищеності, що в свою чергу, обумовлює актуальність удосконалення науково обгрунтованих методів і засобів об'єктивного оцінювання кіберзахисту цих систем. Беручи до уваги те, що якість цього процесу залежить від повноти і достовірності інформації про вразливості і загрози системам, перш за все, важливо проаналізувати існуючі методи і засоби збирання і оброблення такої інформації. Слід підкреслити, що інформація про вразлвості СШІ не є достатньо систематизованою і міститься в джерелах різного типу, а саме базах даних вразливостей, наукових статтях, технічних звітах, що потребує використання сучасних технологій Big Data. Тому актуальними є дослідження і розробки науково-технологічного інструментарію для збору, аналізу і представлення в зручному вигляді даних про вразливості СШІ.

Аналіз інформаційних джерел. У роботі [2] розроблена метамодель, яка складається з компонентів атак, методів та інструментів пом’якшення наслідків. Модель наочно описує зв’язки понять за напрямом кібербезпеки СШІ, але не є повною. Дослідження оцінки загроз у машинному навчанні [3] є певним енциклопедичним документом, але йому не вистачає деталізації щодо СШІ, зокрема, використання ШІ як сервісу. Отже, з огляду на аналіз [1-3] та інших публікацій і розробок об’єктивним є висновок про відсутність досконалих концептуальних моделей і методів збору інформації про вразливості для оцінювання кібербезпеки СШІ.

Метою досліджень є розроблення методу опису процесів збору і аналізу вразливостей СШІ на базі відомої моделі IDEF (Integrated Definition).

Задачі роботи полягають у проведенні критичного аналізу існуючих методів та засобів збору інформації про вразливості для оцінювання кібербезпеки СШІ, формуванні підходу до аналізу, розробленні опису процесу збору та аналізу вразливостей СШІ за допомогою моделі IDEF, ілюстрації особливостей цієї моделі для системи NLP як сервісу та обґрунтуванню напрямів майбтніх досліджень.

Підхід. Пропоновний підхід базується на систематизованому зборі, обробці, нормалізації та поєднанні розрізненої інформації з різних сховищ даних, різних за джерелами наповнення, змістом і форматом, для отримання актуальної, структурированої інформації про вразливості СШІ. Функційний модель цих процесів базується на IDEF нотації.

Для реалізації будемо використовувати програмні засоби: Apache Spark, NumPy, Pandas, SciPy, Matplotlib, Scikit-learn, spaCy та BeautifulSoup.

Висновки. В даній роботі зроблений перший крок до подолання проблеми пошуку та представлення інформації про вразливості СШІ, а саме розроблена трирівнева IDEF модель процесу збору та аналізу інформації про вразливості СШІ для оцінювання кібербезпеки з використанням інструментів великих даних. Визначено показники якості, які оцінюють повноту, узгодженість і достовірність інформації. Подальші дослідження будуть присвячені розробленню методів аналізу і оновлення інформації про вразливості компонентів і СШІ як сервісу та визначення їх критичності, а а також оцінювання та забезпечення кібербезпеки СШІ шляхом аналізу наслідків атак на вразливості та вибору контрзаходів.

Список літератури

1. Неретін, О., Харченко, В. (2022). Забезпечення кібербезпеки систем штучного інтелекту: аналіз вразливостей, атак і контрзаходів. Вісник Національного університету «Львівська політехніка». Інформаційні системи та мережі. Випуск 12, 2022.

2. Fazelnia, M., Khokhlov, I., Mirakhorli, M. (2022). Attacks, Defenses, And Tools: A Framework To Facilitate Robust AI/ML Systems [Online]. Available at: https://arxiv.org/abs/2202.09465.

3. Tidjon, L.N., Khomh, F. (2022). Threat Assessment in Machine Learning based Systems [Online]. Available at: https://arxiv.org/abs/2207.00091.

Відомості про авторів

Неретін Олексій Сергійович, аспірант кафедри комп'ютерних систем, мереж і кібербезпеки, м.т. 099-367-00-71, o.s.neretin@csn.khai.edu

Харченко Вячеслав Сергійович, завідувач кафедри комп'ютерних систем, мереж і кібербезпеки, д.т.н., професор, v.kharchenko@csn.khai.edu