UDC 004.056:004.8

DETECTION AND CLASSIFICATION OF NETWORK INTRUSIONS USING ENSEMBLE LEARNING METHODS

Ivchenko. V.V.
valera2071w@gmail.com
Cherkasy State Business College
Ivanova. I.V.
Cherkasy, Ukraine

The presented work is dedicated to the development of intelligent systems for identifying cyber threats and network anomalies. This study was implemented within the framework of a specialized English-language Python programming course. A feature of the training was the full integration of technical English in the educational process: from studying the documentation of machine learning libraries to communicating with teachers and colleagues exclusively in English for discussing model architectures and experimental results.

The rapid development of cloud technologies and network infrastructure allows for the processing of vast amounts of traffic, which significantly increases the need for robust security monitoring in corporate and state systems. At the same time, the increasing complexity of cyber-attacks creates new challenges in the field of information processing and real-time pattern recognition.

In the conditions of work on the project «Network Intrusion Detection» based on the Kaggle platform, an analysis was conducted of data sequences representing network packet attributes. The main task was to predict different types of network behavior: normal traffic and various categories of malicious activities, including DoS attacks and unauthorized access attempts.

Among the main factors that provided high precision in the developed system, the following can be highlighted:

- Using the XGBoost gradient boosting algorithm, which demonstrated high efficiency in processing non-linear dependencies in network data.

- Thorough preliminary data processing, which included normalization of numerical features, encoding of categorical variables, and handling of imbalanced classes.
- Feature engineering aimed at identifying the most significant indicators of anomalous behavior in traffic flows.
- Application of comprehensive evaluation metrics (Precision, Recall, F1-score) for a deep analysis of model performance in security-critical conditions.

An important role in the implementation of the project was played by proficiency in technical English, which allowed for the effective use of global resources, such as the Kaggle platform, and the integration of modern machine learning techniques into the development. According to the results of testing, a high classification accuracy (0.94) was achieved, which confirms the efficiency of the chosen engineering approach.

In this way, it can be concluded that the combination of Python development skills with fluency in English in a professional environment is a key factor in the successful implementation of complex IT security projects. Effective threat detection requires a comprehensive approach that includes using modern machine learning technologies and constant improvement of a specialist's professional competencies.

References:

1. Kaggle. Network Intrusion Detection Dataset. Access mode: kaggle.com.
2. XGBoost Developers. XGBoost Documentation. Access mode: readthedocs.io.
3. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD, 2016. P. 785–794.
4. NIST Special Publication 800-94. Guide to Intrusion Detection and Prevention Systems. Gaithersburg, 2021.