Штучний інтелект у системах виявлення і запобігання кібератакам: перспективи та виклики

Денис Котенко[1], Юрій Хлапонін[2]

[1,2] Київський національний університет будівництва і архітектури
пр-т Повітрофлотський, 31, Київ, Україна, 03680
[1] kote.denys@gmail.com, https://orcid.org/0009-0005-5888-2143
[2] y.khlaponin@gmail.com, https://orcid.org/0000-0002-9287-0817

Received 13.05.2024, accepted 20.05.2024
https://doi.org/10.32347/uwt.2024.14.1203

Анотація. Стаття розглядає роль штучного інтелекту (ШІ) у системах виявлення і запобігання кібератакам. Автори аналізують поточний стан досліджень та розвитку технологій в цій області, а також визначають перспективи і виклики, з якими стикаються дослідники та практики. Стаття досліджує різноманітні методи та підходи до використання ШІ для виявлення та запобігання кібератакам, включаючи машинне навчання, аналіз поведінки, техніки з обробки природної мови та інші. Також наведений приклад побудови нейронної мережі для відслідковування аномалій.

Ключові слова: штучний інтелект, кібербезпека, виявлення кібератак, запобігання кібератакам, машинне навчання, глибоке навчання, аналіз великих даних, нейронні мережі.

Денис Котенко
магістр Національного авіаційного університету, факультет кібербезпеки та програмної інженерії, спеціаліст: 121 «Інженерія програмного забезпечення»

Юрій Хлапонін
завідувач кафедри кібербезпеки та комп'ютерної інженерії
д.т.н., професор, академік Української академії наук

ПОСТАНОВКА ПРОБЛЕМИ

В сучасному цифровому світі кібербезпека стала надзвичайно важливою проблемою, яка безпосередньо впливає на безпеку, приватність та економічну стабільність індивідів, компаній та навіть держав. Зростання кількості та складності кіберзагроз, таких як хакерські атаки, витоки даних, фішингові кампанії та інші форми кіберзлочинності, свідчить про необхідність посилення заходів з кібербезпеки.

Практично всі сфери життя, включаючи фінанси, медицину, транспорт, енергетику та виробництво, відчутно залежать від інформаційних технологій. Тому навіть незначне порушення кібербезпеки може призвести до серйозних наслідків, таких як втрата конфіденційної інформації, фінансові збитки або навіть загроза життю.

Розвиток глобальних мереж, зростання кількості підключених пристроїв до Інтернету речей (IoT)[1], а також розвиток штучного інтелекту, з одного боку, сприяє збільшенню ефективності та розвитку суспільства, а з іншого - створює нові потенційні точки вразливості, які можуть бути використані зловмисниками для здійснення кібератак.

Сучасний ландшафт кіберзлочинності характеризується постійним зростанням обсягів та складності кібератак. Зловмисники постійно вдосконалюють свої методи та тактики, щоб обійти захист інформаційних систем. У такому контексті поява ефективних засобів виявлення та запобігання кібератакам стає надзвичайно важливою.

По-перше, потрібно враховувати швидкість, з якою здійснюються кібератаки. Зловмисники можуть впроваджувати нові техніки атаки в режимі реального часу, і для

реагування на ці загрози необхідно мати системи, які здатні виявляти та реагувати на них негайно.

По-друге, важливо розуміти, що багато з кібератак можуть бути досить складними та хитрими, і традиційні методи виявлення можуть бути неефективними. Тому потрібні інтелектуальні системи, які здатні аналізувати великі обсяги даних, виявляти аномалії та патерни, які можуть свідчити про потенційні загрози.

Отже, розробка та впровадження ефективних засобів виявлення і запобігання кібератакам стає критично важливим завданням для забезпечення безпеки та стабільності в цифровому світі. Штучний інтелект в цьому контексті може виявитися потужним інструментом, сприяючи удосконаленню систем кібербезпеки та зменшенню вразливості інформаційних систем до кібератак.

МЕТА СТАТТІ

Метою статті є дослідження ролі ШІ у системах виявлення та запобігання кібератакам. Автори прагнуть проаналізувати поточний стан досліджень і технологій у цій галузі, визначити перспективи та виклики, що стоять перед дослідниками і практиками, а також оцінити різноманітні методи та підходи до застосування ШІ для кібербезпеки. Стаття також надає практичний приклад побудови нейронної мережі для відслідковування аномалій, демонструючи реальні можливості та переваги використання ШІ у боротьбі з кіберзагрозами.

ПЕРЕВАГИ ВИКОРИСТАННЯ ШІ В СИСТЕМАХ КІБЕРБЕЗПЕКИ

ШІ може значно полегшити та прискорити процес виявлення потенційних кіберзагроз за допомогою автоматизації. Основні переваги автоматизації виявлення загроз включають:

- Швидкість і ефективність: ШІ здатний аналізувати великі обсяги даних за короткий період часу, що дозволяє виявляти потенційні загрози негайно після їх виникнення.
- Навчання на основі даних: Системи ШІ можуть навчатися на основі історичних даних про кібератаки, а також на основі нових, актуальних вхідних даних, що дозволяє постійно удосконалювати їхню ефективність у виявленні загроз.
- Виявлення невидимих патернів: ШІ здатний виявляти навіть тонкі аномалії та патерни, які можуть бути непомітними для традиційних методів виявлення загроз.
- Мінімізація людського втручання: Автоматизовані системи виявлення загроз можуть значно зменшити потребу в людському втручанні, оскільки вони здатні реагувати на загрози без прямого контролю оператора.
- Скорочення часу реакції: Завдяки швидкій реакції на загрози, автоматизовані системи можуть допомогти у мінімізації часу між виявленням загрози та введенням в дію заходів щодо її запобігання або усунення [2].

Однією з ключових переваг використання ШІ в кібербезпеці є його здатність аналізувати великі обсяги даних для виявлення патернів та аномалій, які можуть свідчити про кіберзагрози. Ця здатність стає все більш важливою, оскільки обсяг даних, які генеруються та збираються, постійно зростає. Людські аналітики просто не в змозі впоратися з таким обсягом даних самостійно, тому ШІ є незамінним інструментом для виявлення кіберзагроз.

ОБМЕЖЕННЯ ТА ВИКЛИКИ

Одним із ключових викликів у використанні ШІ в системах виявлення і запобігання кібератакам є обмеженість якості вихідних даних. Більшість алгоритмів машинного навчання потребують великої кількості якісних даних для ефективної роботи. Проте, в контексті кібербезпеки, дані можуть бути обмеженими через обмежену доступність, недостатню репрезентативність або навіть змінність в часі. Недоліки у якості даних можуть призводити до помилкових виявлень або недооцінки загроз.

Ще одним важливим фактором є вартість

розробки та впровадження. Створення і підтримка таких систем може вимагати значних витрат на інфраструктуру, спеціалізовані кадри та постійне оновлення технологій. Крім того, доступність високоякісних алгоритмів та інструментів для маленьких або середніх компаній може бути обмеженою, що робить їх вразливими перед кіберзагрозами.

Іншою проблемою, з якою зіштовхуються системи виявлення і запобігання кібератакам, є проблема хибнопозитивних та хибнонегативних результатів. Хибнопозитивні результати ведуть до надмірного визначення загроз, що може призвести до невиправданих витрат ресурсів та обмежень для користувачів. З іншого боку, хибнонегативні результати можуть пропускати справжні загрози, що може призвести до серйозних наслідків для безпеки даних.

Основною вимогою для ефективної роботи систем виявлення і запобігання кібератакам на основі ШІ є необхідність постійного навчання та оновлення моделей. Кіберзлочинці постійно змінюють свої тактики, щоб уникнути виявлення, тому системи повинні бути готові адаптуватися до нових загроз. Це може вимагати постійного моніторингу, аналізу та підтримки з боку експертів з кібербезпеки.

ПЕРСПЕКТИВИ ВПРОВАДЖЕННЯ СИСТЕМ ШІ ДЛЯ КІБЕРБЕЗПЕКИ

Здатність до аналізу великих обсягів даних для виявлення патернів та аномалій за допомогою ШІ в кібербезпеці можливо реалізувати за допомогою різноманітних засобів, включаючи:

1. Методи машинного навчання: Використання алгоритмів машинного навчання, таких як нейронні мережі, дерева рішень, алгоритми кластеризації тощо, для автоматизованого аналізу та ідентифікації патернів великих обсягів даних у реальному часі [3].
2. Аналіз поведінки користувачів: Використання ШІ для аналізу поведінки користувачів та виявлення аномальних дій або змін у звичних патернах, що можуть свідчити про потенційні кібератаки або порушення безпеки.
3. Системи виявлення вторгнень (IDS): Використання систем виявлення вторгнень, які базуються на штучному інтелекті, для пошуку аномалій у мережевому трафіку або системних даних, що може свідчити про кіберзагрози [4].
4. Аналіз журналів подій (SIEM): Використання систем управління інформаційною безпекою та подій (SIEM) з інтегрованими засобами аналізу даних для виявлення незвичайних або підозрілих активностей у великих обсягах журналів подій [4].
5. Технології Big Data: Використання технологій обробки великих обсягів даних (Big Data) для збору, зберігання та аналізу великого обсягу даних у реальному часі з метою виявлення патернів та аномалій [5].

Ці засоби дозволяють створювати ефективні системи виявлення та запобігання кібератакам, засновані на здатності ШІ до аналізу великих обсягів даних та виявлення аномалій.

Самого ШІ недостатньо для відслідковування загроз. Тому його вбудовують в інші комп’ютерні системи та навчають на існуючих даних, такі системи називають Інтегровані системи (ІС). ІС, які поєднують у собі ШІ з іншими методами кіберзахисту, стають все більш важливим елементом у сфері інформаційної безпеки. Ці системи забезпечують високий рівень захисту від кіберзагроз, використовуючи технології машинного навчання, глибокого аналізу даних та автоматизації процесів.

Будова та архітектура інтегрованих систем:

- Сенсорний шар (Системи Моніторингу): Інтегровані системи зазвичай починають зі збору даних з різноманітних джерел, таких як мережеві логи, системні журнали, вихідні файли антивірусного програмного забезпечення тощо. Системи моніторингу можуть виявляти аномальну активність, потенційні загрози та ненормальні патерни в поведінці систем.
- Аналітичний шар (Машинне Навчання та

Аналіз Даних): Дані, зібрані на сенсорному рівні, обробляються та аналізуються з використанням алгоритмів машинного навчання, що дозволяє виявляти відхилення від звичайних патернів, а також класифікувати їх як загрози або безпечні події.

- Системи Відповіді (Відслідковування та Реагування): Після виявлення потенційних загроз системи можуть автоматично або за допомогою оператора запускати процеси відслідковування та реагування. Це може включати блокування підозрілих IP-адрес, ізоляцію компрометованих систем або автоматичне оновлення політик безпеки.
- Системи Керування та Моніторингу: Ці системи відповідають за керування та моніторинг інтегрованою платформою кіберзахисту в цілому. Вони забезпечують операторам можливість керувати параметрами системи, отримувати повідомлення про події та стан безпеки, а також вносити необхідні корективи в діючі політики безпеки.

Приклади інтегрованих систем:

1. IBM QRadar поєднує в собі аналіз журналів, потоків трафіку та алгоритми машинного навчання для виявлення кіберзагроз. Система автоматично аналізує дані, виявляє аномалії та специфічні патерни, які можуть свідчити про атаки.
2. Splunk Enterprise Security використовує технології ШІ для виявлення загроз та аналізу даних. Його машинне навчання допомагає виявляти аномальну активність та ідентифікувати загрози в реальному часі.
3. Darktrace використовує алгоритми нейронних мереж для пошуку та виявлення аномальної поведінки в мережі. Система навчається на основі нормальних патернів та може виявляти відхилення від цих патернів, що може свідчити про кібератаку.
4. CyberArk є передовою платформою для забезпечення кібербезпеки, яка спеціалізується на управлінні привілейованими доступами та захисті критично важливих облікових даних. Використання ШІ у CyberArk надає можливість значно покращити захист від внутрішніх та зовнішніх загроз, забезпечуючи високу ефективність та автоматизацію процесів виявлення і реагування на кібератаки.

СТВОРЕННЯ ПРОСТОЇ НЕЙРОННОЇ МОДЕЛІ ДЛЯ ВИЯВЛЕННЯ АНОМАЛІЙ

Аномалії в даних можуть бути ознакою ненормальної поведінки або помилок у даних. Нейронні мережі можуть бути використані для виявлення аномалій шляхом навчання на нормальних даних і виявлення відмінностей в нових даних.

Далі наведений простий приклад виявлення аномалії за допомогою базового автокодувальника, створеного за допомогою мови Python та фреймворку TensorFlow і Keras. Автокодувальник (autoencoder) — це тип нейронної мережі, яка вчиться стискати дані, а потім реконструювати їх. Такі нейронні мережі складаються з кодера (стискає вхідні дані в представлення нижчої розмірності) і декодера (відновлює вихідні дані зі стисненого представлення). Виявлення аномалій за допомогою автокодерів базується на припущенні, що аномалії матимуть більше помилок при реконструкції [6]. Аномалії можна виявити, вимірявши різницю між початковим входом і реконструйованим виходом.

Далі наведений код, який можна виконати в середовищі Python чи Google Codelab:

```
# Імпортувати потрібні бібліотеки:
import numpy as np
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
import matplotlib.pyplot as plt

# Створити функцію генерації нормальних та аномальних даних:
def generate_data():
    # Згенерувати нормальні дані
    normal_data = np.random.normal(loc=0, scale=0.1, size=(1000, 10))

    # Створити аномалії
    anomalies = np.random.normal(loc=2, scale=1, size=(20, 10))

    # Скомбінувати аномалії зі звичайними даними
    data = np.vstack([normal_data, anomalies])
    np.random.shuffle(data)

    # Нормалізувати дані
    data = (data - np.mean(data)) / np.std(data)
    return data

# Створити функцію візуалізації даних:
def plot_data(data, anomalies_indices=[]):
    plt.figure(figsize=(10, 6))
    plt.scatter(data[:, 0], data[:, 1], label='Normal Data', color='blue')
    if len(anomalies_indices) > 0:
        plt.scatter(data[anomalies_indices, 0], data[anomalies_indices, 1],
                    label='Anomalies', color='red', marker='x')
    plt.title('Anomaly Detection with Autoencoder')
    plt.xlabel('Feature 1')
    plt.ylabel('Feature 2')
    plt.legend()
    plt.grid(True)
    plt.show()

# Візуалізація створених даних
# Створення даних
data = generate_data()

# Відображення даних
plot_data(data)
```

Рис. 1. Програмний код для генерації тестових нормальних та аномальних даних, а також їх відображення (на мові Python)

Fig. 1. Program code for generating test normal and abnormal data, as well as their display (in Python)

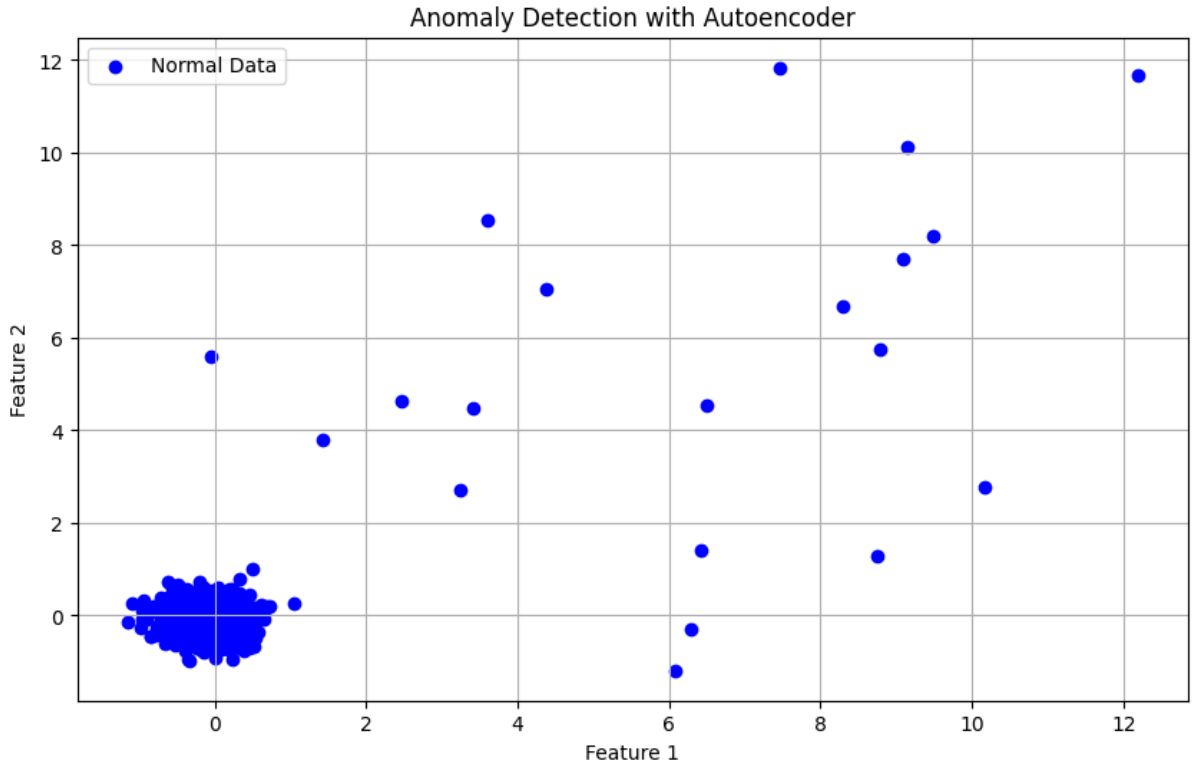

Рис. 2. Випадково згенеровані дані з аномальними відхиленнями

Fig. 2. Randomly generated data with anomalous deviations

Як бачимо, нормальні дані знаходяться в межах 0-1 по вісі абсцис та ординат, тоді як аномальні дані розміщуються в межах 2-12 по вісі абсцис та ординат.

```
# Створити модель автокодувальника
input_dim = data.shape[1]

# Кодувальник
encoder_input = keras.Input(shape=(input_dim,))
encoder_output = layers.Dense(64, activation='relu')(encoder_input)
encoder_output = layers.Dense(32, activation='relu')(encoder_output)

# Декодувальник
decoder_output = layers.Dense(64, activation='relu')(encoder_output)
decoder_output = layers.Dense(input_dim, activation='sigmoid')(decoder_output)

# Автокодувальник
autoencoder = keras.Model(encoder_input, decoder_output)
autoencoder.compile(optimizer='adam', loss='mse')
autoencoder.summary()

# Тренування моделі
autoencoder.fit(data, data, epochs=50, batch_size=32)
```

Рис. 3. Програмування нейронної моделі автокодувальника та його подальше навчання

Fig. 3. Programming the neural model of the autoencoder and its further training

```
# Реконструкція вхідних даних
reconstructed_data = autoencoder.predict(data)

# Підрахування кількості помилок
mse = np.mean(np.square(data - reconstructed_data), axis=1)

# Встановити граничну межу для визначення аномалій
threshold = np.mean(mse) + 2 * np.std(mse)

# Ідентифікація аномалій
anomalies_indices = np.where(mse > threshold)[0]

print("Indices of anomalies:", anomalies_indices)

# Відображення знайдених аномалій
plot_data(data, anomalies_indices)
```

Рис. 4. Визначення аномалій за допомогою натренованої нейронної мережі

Fig. 4. Determining anomalies using a trained neural network

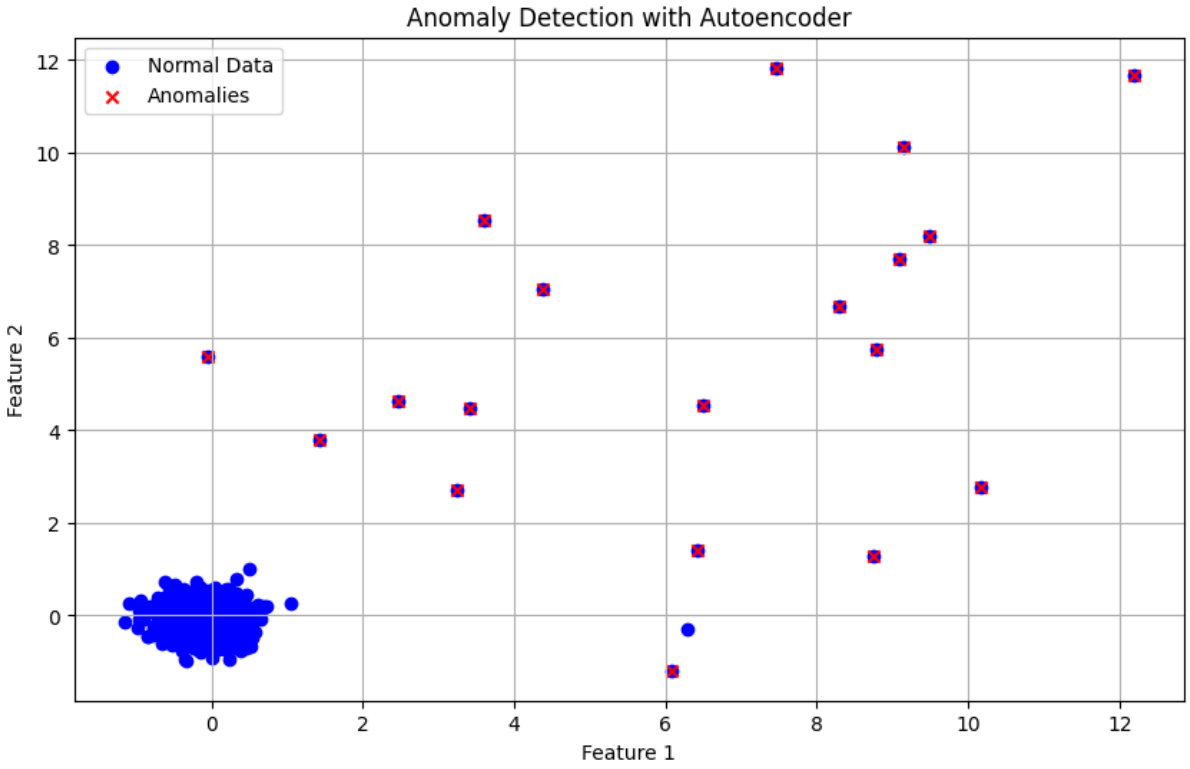

Рис. 5. Визначення аномалій за допомогою натренованої нейронної мережі

Fig. 5. Determining anomalies using a trained neural network

Як бачимо, в результаті створеної нами нейронної моделі, ми в змозі відрізнити аномальні дані від звичайних. Також бачимо, що існує хибнопозитивна похибка, при якій не вдалось розпізнати одну аномалію.

ВИСНОВКИ

ШІ має значний потенціал для покращення кібербезпеки. Його здатність аналізувати великі обсяги даних, виявляти патерни та аномалії може значно покращити

ефективність систем виявлення та запобігання кібератакам.

Однак важливо пам'ятати про обмеження та виклики використання ШІ в кібербезпеці. Для успішного впровадження систем ШІ для кібербезпеки необхідно вирішити такі проблеми, як обмеженість якості даних, висока вартість розробки та впровадження, а також ризик хибнопозитивних та хибнонегативних результатів.

Також ми розглянули приклад створення простої нейронної мережі для виявлення аномалій та переконались, що такі мережі не можуть надавати 100-відсоткову гарантію в правильності своєї роботи, тим не менш, в змозі суттєво автоматизувати процес пошуку та ідентифікації відхилень від норми.

В цілому, ШІ є потужним інструментом, який може допомогти покращити кібербезпеку та захистити інформаційні системи від кіберзагроз.

ЛІТЕРАТУРА

1. **Власенко М., Хлапонін Ю**. (2024). Інтернет речей (IoT) у світовій практиці: огляд та аналіз. Pidvodni Tehnologii, (13), 21–27. https://doi.org/10.32347/uwt.2023.13.1202
2. **Kabbas A., Alharthi A., Munshi A.** Artificial Intelligence Applications in Cybersecurity. IJCSNS International Journal of Computer Science and Network Security, VOL.20 No.2, Fabruary 2020. 120-124. https://ru.scribd.com/document/487200875/research-paper-5
3. **Al Musawi, Ahmad.** (2018). Introduction to Machine Learning. https://www.researchgate.net/publication/323108787_Introduction_to_Machine_Learning
4. What is an intrusion detection system (IDS)? https://www.ibm.com/topics/intrusion-detection-system
5. What is Big Data Analytics for Cyber Security https://www.sangfor.com/blog/cybersecurity/what-is-big-data-analytics-for-cyber-security
6. Anomaly detection with Keras, TensorFlow, and Deep Learning. March 2020 https://pyimagesearch.com/2020/03/02/anomaly-detection-with-keras-tensorflow-and-deep-learning/

Artificial intelligence in cyber attack detection and prevention systems: prospects and challenges

Denis Kotenko, Yuri Khlaponin

Abstract. The article examines the role of artificial intelligence (AI) in cyber attack detection and prevention systems. The authors analyze the current state of research and technology development in this area, as well as identify prospects and challenges facing researchers and practitioners. The article explores a variety of techniques and approaches to using AI to detect and prevent cyberattacks, including machine learning, behavioral analysis, natural language processing techniques, and more. An example of building a neural network for tracking anomalies is also given.

Keywords: artificial intelligence, cyber security, cyber attack detection, cyber attack prevention, machine learning, deep learning, big data analysis, neural networks.