ხელოვნურ ინტელექტზე დაფუძნებული ინფორმაციულ სისტემებში შეჭრის აღმოჩენის სისტემები

1. შესავალი

დღევანდელ სწრაფად განვითარებად ციფრულ გარემოში, კიბერ საფრთხეების სიხშირე და სირთულე გეომეტრიული პროგრესიით იზრდება. ტრადიციული IDS მეთოდები, რომლებიც ხშირად ეყრდნობიან წინასწარ განსაზღვრული წესების კომპლექტს და ხელმოწერებზე დაფუძნებულ გამოვლენის ტექნიკას, იბრძვიან შენარჩუნებული თავდასხმის ვექტორებთან. შედეგად, იზრდება ინტელექტუალური, ადაპტირებული სისტემების საჭიროება, რომელსაც შეუძლია აღმოაჩინოს უცნობი და ნულოვანი დღის საფრთხეები. მანქანური დასწავლა (ML) გთავაზობთ ამ პრობლემის ინოვაციურ გადაწყვეტას, რაც IDS-ს საშუალებას აძლევს მუდმივად ისწავლოს მონაცემებიდან, გაუმჯობესდეს დროთა განმავლობაში და ეფექტურად უპასუხოს უსაფრთხოების ახალ გამოწვევებს.

სემინარზე განხილული იქნება AI-ზე ორიენტირებული შეჭრის აღმოჩენის სისტემების (AI-IDS) როლი თანამედროვე კიბერუსაფრთხოებაში. კვლევებით დასაბუთებულია, თუ როგორ გამოიყენება მანქანური დასწავლის ალგორითმები - განსაკუთრებით ღრმა დასწავლის, განმტკიცებით დასწავლა და უკონტროლო დასწავლა - IDSის ეფექტური ფუნქციონირებისათვის. წარმოდგენილი მიდგომა, აუმჯობესებს გამოვლენის მაჩვენებელს, ამცირებს ცრუ პოზიტიურ და უზრუნველყოფს პროაქტიულ მიდგომას უსაფრთხოებასთან დაკავშირებით. გარდა ამისა, Python-ის გამოყენებით პრაქტიკული განხორციელება მოცემულია იმის ფაქტების საილუსტრაციოდ, თუ როგორ შეიძლება მანქანური დასწავლის გამოყენება ინფორმაციულ სისტემებში შეჭრის გამოვლენის გამოწვევების გადასაჭრელად.

2. ტრადიციული შეჭრის აღმოჩენის სისტემები

ტრადიციული IDS მეთოდები, როგორიცაა ხელმოწერებზე დაფუძნებული გამოვლენა და ანომალიებზე დაფუძნებული გამოვლენა, ფართოდ გამოიყენება ათწლეულების განმავლობაში. ხელმოწერებზე დაფუძნებული სისტემები ემთხვევა ქსელის ტრაფიკს შეტევის ცნობილი ნიმუშების მონაცემთა ბაზას, ხოლო ანომალიებზე დაფუძნებული სისტემები იდენტიფიცირებენ გადახრებს ნორმალური ქცევისგან. მიუხედავად იმისა, რომ ეფექტურია ცნობილი თავდასხმებისთვის, ეს სისტემები ხშირად არასაკმარისია ახალი ან მოწინავე მდგრადი საფრთხეების (APTs) გამოსავლენად, რადგან ისინი დიდწილად ეყრდნობიან წინასწარ განსაზღვრულ შაბლონებს და ზღურბლებს.

3. მანქანური დასწავლის ზრდა IDS-ში

მანქანური დასწავლის თანამედროვე მიდგომებს შემოაქვს პარადიგმის ცვლილება შეჭრის გამოვლენის სფეროში. სტატისტიკური მოდელებისა და ალგორითმების გამოყენებით, ML სისტემებს შეუძლიათ ავტომატურად ისწავლონ შაბლონები დიდი მონაცემთა ნაკრებიდან და განახორციელონ პროგნოზები აშკარა პროგრამირების გარეშე. IDS-ის კონტექსტში, ML ალგორითმებს შეუძლიათ:

• *უცნობი თავდასხმების გამოვლენა:* ქსელის ნორმალური ქცევიდან სწავლით, ML სისტემებს შეუძლიათ დაადგინონ ანომალიები, რომლებიც გადახრის ტიპიურ შაბლონებს, მაშინაც კი, თუ ეს ანომალიები ადრე უცნობია.

• *აღმოჩენის სიზუსტის გაუმჯობესება:* საკმარისი სასწავლო მონაცემებით, ML ალგორითმებს შეუძლიათ დახვეწონ თავიანთი მოდელები დროთა განმავლობაში, რაც აუმჯობესებს მათ უნარს განასხვავოს კეთილთვისებიანი და მავნე მოქმედებები.

• ადაპტირება განვითარებულ საფრთხეებთან: ML ალგორითმებს შეუძლიათ მუდმივად ისწავლონ და მოერგოს ახალ მონაცემებს, რაც IDS-ს საშუალებას აძლევს დარჩეს კიბერ საფრთხეების განვითარებად ბუნებასთან.

მათემატიკურად, ანომალიის გამოვლენის პროცესი შეიძლება განიხილებოდეს, როგორც გადახრების იდენტიფიცირება ნასწავლი განაწილებიდან სადაც **p(x)p(x)p(x)** წარმოადგენს **სავარაუდო (ნორმალური) მონაცემთა განაწილების ალბათობის სიმკვრივის ფუნქციას**. ეს ნიშნავს, რომ:

- გვაქვს მონაცემთა დიდი ნაკრები (მაგ., ქსელის ტრაფიკის თვისებები: პაკეტის ზომა, სიხშირე, დროის ინტერვალი და ა.შ.).
- ამ მონაცემებისგან ვსწავლობთ (შევაფასებთ) განაწილებას – მაგალითად, ნორმალური განაწილება, Gaussian mixture models ან სხვა არასპეციფიკური განაწილება.
- შედეგად ვიღებთ ფუნქციას p(x)p(x)p(x), რომელიც თითოეულ ახალ მონაცემს xxx – ანუ ქსელის ერთ კონკრეტულ პაკეტს ან ტრაფიკის მდგომარეობას – ანიჭებს ალბათობას, რომ ის „ნორმალურია“.

მაგალითად: თუ p(x) = 0.002, ეს ნიშნავს, რომ მოცემული ტრაფიკი ძალიან არაბუნებრივია ჩვეულებრივი (ნასწავლი) მონაცემების მიმართ.

L(x)=−logp(x)L(x) = -\log p(x)L(x)=−logp(x) – ეს არის **ნეგატიური ლოგარითმული ალბათობა**, ანუ მონაცემის „სიურპრიზულობის“ ან **ინფორმაციული შინაარსის** რაოდენობა.

- რაც უფრო დაბალია p(x), ანუ რაც უფრო უცნაურია ეს მონაცემი, მით უფრო მაღალია L(x).
- ეს იდეა ეფუძნება **ინფორმაციის თეორიას**: იშვიათი მოვლენები მეტ ინფორმაციას შეიცავს.

მაგ. თუ p(x) = 0.01 → L(x) = −log(0.01) ≈ 4.6

ხოლო თუ p(x) = 0.9 → L(x) = −log(0.9) ≈ 0.1

მიდგომის მიზანია: **განსაზღვროს ზღვარი T**, რომელიც განსაზღვრავს რა დონემდე შეიძლება გაიზარდოს L(x)L(x)L(x) ისე, რომ მონაცემი ჯერ კიდევ ჩაითვალოს „ნორმალურად“.

თუ L(x)>TL(x) > TL(x)>T, ეს ნიშნავს, რომ ეს მონაცემი **ანომალიურია** და შეიძლება იყოს პოტენციური კიბერშეჭრის მაჩვენებელი. აღწერილი პირობის **პრაქტიკული ინტერპრეტაცია ასეთია, p(x)** – განსაზღვრავს, რამდენად „სავარაუდოა“ მოცემული პაკეტი ჩვენს მიერ ნასწავლ ქსელურ ტრაფიკში. **L(x)** – განსაზღვრავს, რამდენად საეჭვოა ეს პაკეტი: დიდი მნიშვნელობა = პოტენციური ანომალია.

4. მანქანური დასწავლის ალგორითმები ინფორმაციულ სიტემებში შეჭრის აღმოჩენისთვის

მანქანური დასწავლის მეთოდებმა აჩვენა დადებითი შედეგი IDS-ის გაძლიერებაში. მათ შორის ყველაზე გამორჩეულია:

• **ღრმა სწავლება (DL):** ღრმა ნეირონული ქსელები (DNN), განსაკუთრებით კონვოლუციური ნეირონული ქსელები (CNN) და განმეორებადი ნეირონული ქსელები (RNNs), ეფექტურია ქსელის ტრაფიკის რთული შაბლონების მოდელირებასა და ანალიზში. მათემატიკურად ღრმა დასწავლის მოდელები შეიძლება წარმოდგენილი იყოს შემდეგი სახით:

y=f(x;θ) y = f(x; \theta) y=f(x;θ)

სადაც y არის პროგნოზირებული გამომავალი (მაგ. ნორმალური ან თავდასხმა), x წარმოადგენს შეყვანის მახასიათებლებს (ქსელის მონაცემებს) და θ\thetaθ წარმოადგენს მოდელის პარამეტრებს (წონებს). მოდელი მომზადებულია ზარალის ფუნქციის L მინიმიზაციის გზით, როგორიცაა ჯვარედინი ენტროპიის დანაკარგი:

L(θ)=−∑i=1N[yilog⁡(y^i)+(1−yi)log⁡(1−y^i)]L(\theta) = -\sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]L(θ)=−i=1∑N[yilog(y^i)+(1−yi)log(1−y^i)]

სადაც yiy_iyi არის ჭეშმარიტი ეტიკეტები და y^i\hat{y}_iy^i არის პროგნოზირებული ალბათობები.

• **განმტკიცებით დასწავლა (RL):** RL მოდელები იყენებენ აგენტს, რომელიც ურთიერთქმედებს მის გარემოსთან და სწავლობს ოპტიმალურ მოქმედებებს ჯილდოებზე დაყრდნობით. შეჭრის აღმოსაჩენად აგენტს შეეძლო გაეგო, თუ რომელი ქმედებები (მაგ., გარკვეული მოძრაობის შაბლონების მონიშვნა თავდასხმებად) მაქსიმუმს ანიჭებს შეჭრის ზუსტად აღმოჩენის ჯილდოს.

აგენტის ქცევა შეიძლება იყოს მოდელირებული, როგორც π(a|s)\pi(a|s)π(a|s), სადაც aaa არის მოქმედება (მაგ., ტრაფიკის მონიშვნა მავნედ) და sss არის მდგომარეობა (მაგ. ქსელის ტრაფიკის მახასიათებლები). RL აგენტის მიზანია კუმულაციური ჯილდოს მაქსიმიზაცია დროთა განმავლობაში:

R(π)=∑t=0TγtrtR(\pi) = \sum_{t=0}^{T} \gamma^t r_tR(π)=t=0∑Tγtrt

სადაც rtr_trt არის ჯილდო დროის საფეხურზე ttt, და γ\gammaγ არის ფასდაკლების ფაქტორი.

• **ზედამხედველობის გარეშე დასწავლა:** უკონტროლო სწავლის ალგორითმები, როგორიცაა კლასტერირება (მაგ., K-საშუალებები) ან ანომალიების გამოვლენა, სასარგებლოა, როდესაც ეტიკეტირებული თავდასხმის მონაცემები მწირია. ერთი გავრცელებული მეთოდია მონაცემთა წერტილების პოვნა, რომლებიც მნიშვნელოვნად განსხვავდება ქსელის ტრაფიკის უმრავლესობისგან, მანძილის მეტრიკის გამოყენებით, როგორიცაა ევკლიდური მანძილი:

d(x,c)=∑i=1n(xi−ci)2d(x, c) = \sqrt{\sum_{i=1}^{n} (x_i - c_i)^2}d(x,c)=i=1∑n(xi−ci)2

სადაც xxx წარმოადგენს მონაცემთა წერტილს და ccc არის კლასტერის ცენტრი. წერტილები დიდი d(x,c)d(x, c)d(x,c) მნიშვნელობებით ითვლება ანომალიურად.

5. პროგრამული რეალიზაცია/იმპლემენტაცია. Python-ზე დაფუძნებული მანქანური დასწავლის მოდელის დანერგვა შეჭრის აღმოჩენისთვის

ჩვენ გამოვყოფთ - Python-ზე დაფუძნებულ მიდგომას მანქანური სწავლებით აღჭურვილი შეჭრის აღმოჩენის სისტემის დანერგვისთვის პოპულარული ბიბლიოთეკების გამოყენებით, როგორიცაა scikit-learn, pandas და TensorFlow. მიდგომა მოიცავს ქსელის ტრაფიკის მონაცემების წინასწარ დამუშავებას, მანქანური სწავლების მოდელის ტრენინგს და მისი შესრულების შეფასებას.

ნაბიჯი 1: მონაცემთა წინასწარი დამუშავება

პირველი ნაბიჯი არის ქსელის ტრაფიკის მონაცემების ჩატვირთვა და წინასწარი დამუშავება. ამ მაგალითისთვის, ჩვენ ვიყენებთ NSL-KDD მონაცემთა ბაზას, ჩვეულებრივ გამოყენებულ მონაცემთა ბაზას შეჭრის აღმოჩენის ამოცანებისთვის.

python
Copy code

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import StandardScaler

# Load dataset
data = pd.read_csv('KDDTrain+.csv')

# Feature engineering: Convert categorical labels to numeric
label_encoder = LabelEncoder()
data['class'] = label_encoder.fit_transform(data['class'])

# Split data into features (X) and labels (y)
X = data.drop('class', axis=1)
y = data['class']

# Standardize the feature data
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.3, random_state=42)
```

ნაბიჯი 2: მოდელის ტრენინგი და შეფასება

შეჭრის აღმოჩენისთვის, ჩვენ გამოვიყენეთ შემთხვევითი ტყის კლასიფიკატორი, რომელიც ეფექტურია კლასიფიკაციის ამოცანებისთვის. ჩვენ ასევე შევაფასებთ მოდელის სიზუსტეს და შესრულებას ისეთი მეტრიკის გამოყენებით, როგორიცაა სიზუსტე, გახსენება და F1-ქულა.

python
Copy code
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report, accuracy_score

# Initialize the model
model = RandomForestClassifier(n_estimators=100, random_state=42)

# Train the model
model.fit(X_train, y_train)

# Make predictions
y_pred = model.predict(X_test)

# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
report = classification_report(y_test, y_pred)
print(f"Accuracy: {accuracy}")
print("Classification Report:")
print(report)

ნაბიჯი 3: მოდელის გაუმჯობესება და დახვეწა

მოდელის მუშაობის გასაუმჯობესებლად, ჩვენ შეგვიძლია ექსპერიმენტი გავუკეთოთ ჰიპერპარამეტრების რეგულირებას და ჯვარედინი ვალიდაციას.

python
Copy code
from sklearn.model_selection import GridSearchCV

# Hyperparameter tuning for Random Forest
param_grid = {
    'n_estimators': [50, 100, 150],

```
    'max_depth': [10, 20, 30],
    'min_samples_split': [2, 5, 10]
}

grid_search = GridSearchCV(estimator=model, param_grid=param_grid, cv=3, verbose=2, n_jobs=-1)
grid_search.fit(X_train, y_train)

# Best parameters and model evaluation
best_model = grid_search.best_estimator_
y_pred_best = best_model.predict(X_test)

accuracy_best = accuracy_score(y_test, y_pred_best)
report_best = classification_report(y_test, y_pred_best)

print(f"Improved Accuracy: {accuracy_best}")
print("Improved Classification Report:")
print(report_best)
```

6. AI-ზე ორიენტირებული პლატფორმების გამოწვევები სისტემებში შეჭრის გამოვლენაში

AI-ზე ორიენტირებული IDS-ის პოტენციალის მიუხედავად, რამდენიმე გამოწვევა უნდა გადაიჭრას:

• *მონაცემთა ხარისხი და მარკირება:* მაღალი ხარისხის მონაცემები გადამწყვეტია ეფექტური ML მოდელების მომზადებისთვის. ხშირ შემთხვევაში, ეტიკეტირებული თავდასხმის მონაცემების მოპოვება რთულია და ხმაურიანმა ან არასრულმა მონაცემთა ნაკრებებმა შეიძლება გააუარესოს მოდელის შესრულება.

• *მასშტაბურობა:* AI მოდელები, განსაკუთრებით ღრმა სწავლის მოდელები, საჭიროებენ მნიშვნელოვან გამოთვლით რესურსებს. მათი დანერგვა ფართომასშტაბიანი ქსელის გარემოში შეიძლება წარმოქმნას მასშტაბურობის გამოწვევები.

• *ინტერპრეტაცია:* AI-ზე დაფუძნებული IDS-ის ერთ-ერთი კრიტიკული საკითხია შედეგების ინტერპრეტაცია. მიუხედავად იმისა, რომ ღრმა სწავლის მოდელები ძალზე ეფექტურია, მათი „შავი ყუთის“ ბუნება ართულებს იმის გაგებას, თუ რატომ იქნა მიღებული გარკვეული გადაწყვეტილება. ამ გამჭვირვალობის ნაკლებობამ შეიძლება შეაფერხოს სისტემისადმი ნდობა, განსაკუთრებით უსაფრთხოების კრიტიკულ აპლიკაციებში.

7. მომავალი მიმართულებები და პერსპექტივა

AI-ზე ორიენტირებული IDS-ის სფერო ჯერ კიდევ ვითარდება და სამეცნიერო კვლევებში ისახება მისი შესაძლებლობების შემდგომ გაძლიერებას:

• *ჰიბრიდული მიდგომები:* მანქანური სწავლების მრავალი ტექნიკის გაერთიანებამ, როგორიცაა ღრმა სწავლების შერწყმა უკონტროლო სწავლებასთან, შეიძლება გამოიწვიოს უფრო ძლიერი და ადაპტირებული შეჭრის აღმოჩენის სისტემები.

• ფედერაციული სწავლება: კონფიდენციალურობის საკითხების ზრდასთან ერთად, ფედერირებული სწავლამ შეიძლება მისცეს ML მოდელების ტრენინგი დეცენტრალიზებულ მონაცემთა წყაროებში, რაც აუმჯობესებს მონაცემთა უსაფრთხოებას და აძლიერებს გამოვლენის სიზუსტეს.

• *ახსნადი AI (XAI):* ახსნადი AI მოდელების შემუშავებამ შეიძლება მოაგვაროს ინტერპრეტაციის გამოწვევა IDS-ის გადაწყვეტილების მიღების პროცესში მეტი გამჭვირვალობის უზრუნველყოფით, რითაც გაზრდის მომხმარებლის ნდობას.

8. დასკვნა

AI-ზე ორიენტირებული შეჭრის აღმოჩენის სისტემები წარმოადგენს კიბერუსაფრთხოების მომავალს, გვთავაზობს გაფართოებულ შესაძლებლობებს აღმოაჩინოს და უპასუხოს კიბერ საფრთხეების ფართო სპექტრს. მანქანური სწავლის, განსაკუთრებით ღრმა სწავლის, განმამტკიცებელი დასწავლის და ზედამხედველობის გარეშე დასწავლის ინსტრუმენტის გამოყენებით, IDS შეიძლება განუწყვეტლივ განვითარდეს, რათა დახვდეს საფრთხის უფრო დახვეწილ გამოწვევებს.

სემინარში წარმოდგენილი პითონზე დაფუძნებული იმპლემენტაცია გვიჩვენებს, თუ როგორ შეიძლება გამოყენებულ იქნას მანქანური დასწავლა შეჭრის აღმოჩენისთვის, აჩვენებს პრაქტიკულ გადაწყვეტილებებს რეალურ სამყაროში კიბერუსაფრთხოების აპლიკაციებისთვის. დამუშავებული მიდგომა პროგრამისტებს დაეხმარება ისეთი გამოწვევების მოგვარებაში, როგორიცაა მონაცემთა ხარისხი, მასშტაბურობა და ინტერპრეტაცია, გადამწყვეტი იქნება AI-IDS-ის სრული პოტენციალის რეალიზებისთვის.