MASHINALI O‘QITISH USULLARI ASOSIDA FISHING HUJUMLARINI ANIQLASH

Orif Allanov Menglimuratovich
Dotsent, PhD, Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti, Toshkent, O‘zbekiston
allanov@tuit.uz
Hafizov Shukurullo Fayzullo o‘g‘li
magistr, Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti, Toshkent, O‘zbekiston

Annotatsiya: Hozirgi raqamli texnologiyalar rivojlanishi sharoitida fishing hujumlari eng keng tarqalgan kiberxavfsizlik tahdidlaridan biri hisoblanadi. Ushbu hujumlar foydalanuvchilarning shaxsiy ma’lumotlarini, login va parollarini, bank kartalari rekvizitlarini hamda boshqa maxfiy ma’lumotlarni noqonuniy qo‘lga kiritishga qaratilgan bo‘ladi. An’anaviy himoya mexanizmlari zamonaviy fishing hujumlarini aniqlashda har doim ham samarali natija bermaydi. Shu sababli mashinaviy o‘qitish texnologiyalaridan foydalanish dolzarb masalaga aylangan. Mazkur maqolada fishing hujumlarini aniqlashda qo‘llaniladigan mashinaviy o‘qitish usullari, ularning ishlash prinsiplari, afzalliklari va kamchiliklari tahlil qilinadi.
Kalit so‘zlar: fishing hujumi, mashinaviy o‘qitish, kiberxavfsizlik, phishing detection, Random Forest, Neural Network, sun’iy intellekt, axborot xavfsizligi.

Kirish. Internet texnologiyalarining jadal rivojlanishi natijasida axborot almashinuvi, elektron tijorat, masofaviy bank xizmatlari va bulutli texnologiyalar keng qo‘llanila boshlandi. Shu bilan birga kiberjinoyatchilar tomonidan amalga oshiriladigan fishing hujumlari soni ham ortib bormoqda. Fishing hujumlari foydalanuvchilarni soxta veb-saytlar, elektron pochta xabarlari yoki zararli havolalar orqali aldashga asoslangan bo‘lib, ularning asosiy maqsadi maxfiy ma’lumotlarni qo‘lga kiritishdan iborat.

An’anaviy xavfsizlik vositalari, jumladan blacklist va signature-based himoya tizimlari yangi va murakkab fishing hujumlarini aniqlashda yetarli darajada samarali emas. Chunki zamonaviy fishing hujumlari doimiy ravishda o‘zgarib boradi va oddiy filtrlarni chetlab o‘tishga moslashtiriladi. Shu sababli sun’iy intellekt va mashinaviy o‘qitish algoritmlaridan foydalanish fishing hujumlarini avtomatik aniqlashning istiqbolli yechimlaridan biri hisoblanadi.

Mashinaviy o‘qitish usullari katta hajmdagi ma’lumotlarni tahlil qilish, zararli faoliyat belgilarini aniqlash va real vaqt rejimida tahdidlarni aniqlash imkonini beradi.

Ushbu maqolada fishing hujumlarini aniqlash uchun qo‘llaniladigan asosiy mashinaviy o‘qitish algoritmlari tahlil qilinadi hamda ularning samaradorligi o‘rganiladi.

Fishing hujumlari tushunchasi va asosiy xususiyatlari

Fishing hujumi bu foydalanuvchini aldash orqali maxfiy ma’lumotlarni qo‘lga kiritishga qaratilgan kiberhujum turidir. Hujumchi odatda ishonchli tashkilot nomidan soxta elektron pochta, SMS yoki veb-sahifa yaratadi. Foydalanuvchi ushbu resursga ishongan holda login, parol yoki bank ma’lumotlarini kiritadi. Fishing hujumlari bir necha turlarga bo‘linadi:

1.1-jadval. Fishing turi va ularning tavsifi

Fishing turi	Tavsifi
Email phishing	Elektron pochta orqali amalga oshiriladi
Spear phishing	Muayyan shaxs yoki tashkilotga qaratilgan
Smishing	SMS orqali fishing hujumi
Vishing	Telefon qo‘ng‘iroqlari orqali amalga oshiriladi
Clone phishing	Ishonchli xabar nusxasi orqali hujum

Fishing hujumlarining asosiy xavfi shundaki, ular inson omiliga asoslanadi. Ko‘pchilik foydalanuvchilar soxta sahifalarni haqiqiy saytlar bilan adashtiradi.

Mashinaviy o‘qitish usullarining fishing hujumlarini aniqlashdagi roli

Mashinaviy o‘qitish algoritmlari ma’lumotlardan naqshlarni aniqlash va kelajakdagi holatlarni bashorat qilish imkonini beradi. Fishing hujumlarini aniqlashda URL manzillari, domen nomlari, sahifa tarkibi va trafik xususiyatlari tahlil qilinadi.

Fishingni aniqlash jarayoni quyidagi bosqichlardan iborat:

1. Ma’lumotlarni yig‘ish
2. Belgilarni ajratib olish
3. Modelni o‘qitish
4. Testlash va baholash
5. Real vaqt monitoringi

Fishingni aniqlashda quyidagi belgilar muhim hisoblanadi:

- URL uzunligi;
- HTTPS mavjudligi;
- Maxsus belgilar soni;
- Domen yoshi;
- IP manzildan foydalanish;
- Shubhali kalit so‘zlar mavjudligi.

Asosiy mashinaviy o‘qitish algoritmlari

Decision Tree algoritmi

Decision Tree qaror daraxti asosida ishlovchi algoritm bo‘lib, ma’lumotlarni bir nechta shartlar asosida tasniflaydi. Ushbu usul sodda va tushunarli hisoblanadi.

Afzalliklari:

- ishlash tezligi yuqori;
- tushunish oson;
- kichik ma’lumotlarda samarali.

Kamchiliklari:

- overfitting muammosi;
- katta ma’lumotlarda aniqlik pasayishi.

Random Forest algoritmi

Random Forest bir nechta qaror daraxtlari asosida ishlaydi va eng ko‘p ovozga ega natijani tanlaydi. Ushbu algoritm fishing hujumlarini aniqlashda yuqori samaradorlik ko‘rsatadi.

Afzalliklari:

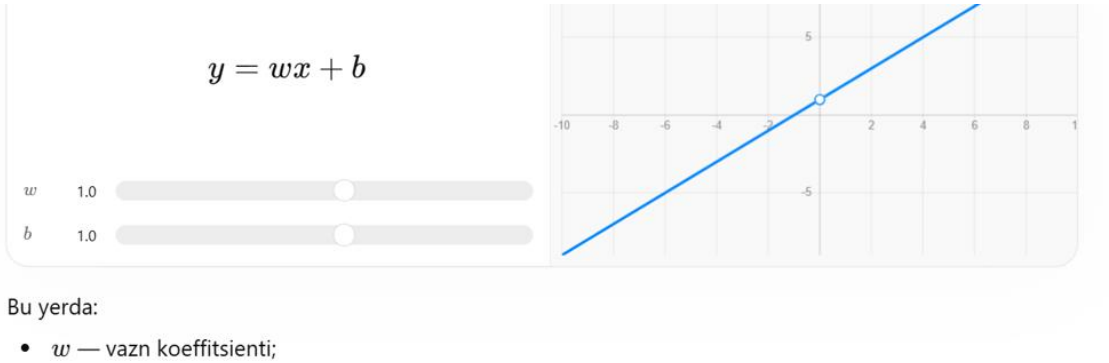
- yuqori aniqlik;
- overfitting ehtimoli past;
- katta hajmdagi ma’lumotlar bilan ishlay oladi.

Kamchiliklari:

- hisoblash xarajati yuqori;
- murakkab tuzilishga ega.

Support Vector Machine (SVM)

SVM algoritmi ma’lumotlarni gipertekislik yordamida ajratadi. Ushbu usul yuqori aniqlik talab qilinadigan vaziyatlarda samarali hisoblanadi.



Bu yerda:

- w — vazn koeffitsienti;
- x — kiruvchi ma’lumot;
- b — siljish parametri.

SVM algoritmi fishing va legitim saytlarni optimal chegaralar yordamida ajratadi.

Neural Network algoritmi

Sun’iy neyron tarmoqlari inson miyasi faoliyatiga o‘xshash tarzda ishlaydi. Ushbu algoritmlar murakkab naqshlarni aniqlashda juda samarali hisoblanadi.

Neural Network afzalliklari:

- yuqori aniqlik;
- katta ma’lumotlar bilan ishlash;
- murakkab hujumlarni aniqlash.

Kamchiliklari:

- katta resurs talab qiladi;
- modelni o‘qitish uzoq davom etadi.

1.2-jadval. Algoritmlar samaradorligini taqqoslash

Algoritm	Aniqlik (%)	Tezlik	Overfitting holati
Decision Tree	89	Yuqori	Yuqori
Random Forest	97	O‘rta	Past

SVM	95	O‘rta	Past
Naive Bayes	87	Yuqori	O‘rta
Neural Network	98	Past	Past

Jadvaldan ko‘rinib turibdiki, Neural Network va Random Forest algoritmlari eng yuqori natijalarni ko‘rsatgan.

Taklif etilayotgan model

Mazkur tadqiqotda fishing hujumlarini aniqlash uchun gibrid model taklif etiladi. Ushbu model URL tahlili, domen tekshiruvi va mashinaviy o‘qitish algoritmlarini birlashtiradi.

Model quyidagi bosqichlarda ishlaydi:

1. URL manzilni qabul qilish
2. Belgilarni ajratish
3. Ma’lumotlarni normalizatsiya qilish
4. Random Forest modeli orqali tahlil qilish
5. Natijani foydalanuvchiga uzatish

Taklif etilgan model real vaqt rejimida fishing saytlarni aniqlash imkonini beradi.

Xulosa

Fishing hujumlari zamonaviy axborot tizimlari uchun jiddiy tahdid hisoblanadi. An’anaviy himoya vositalari yangi turdagi fishing hujumlarini aniqlashda cheklangan imkoniyatlarga ega. Mashinaviy o‘qitish algoritmlaridan foydalanish esa fishing hujumlarini tezkor va aniq aniqlash imkonini beradi.

Tadqiqot davomida turli algoritmlar tahlil qilinib, Random Forest va Neural Network usullari eng samarali natijalarni ko‘rsatgani aniqlandi. Kelgusida chuqur o‘qitish texnologiyalari va real vaqt monitoring tizimlarini integratsiya qilish fishing hujumlariga qarshi kurash samaradorligini yanada oshirishi mumkin.

FOYDALANILGAN ADABIYOTLAR:

1. Ian Goodfellow. Deep Learning. MIT Press, 2016.
2. Christopher Bishop. Pattern Recognition and Machine Learning. Springer, 2006.
3. William Stallings. Network Security Essentials. Pearson, 2017.
4. Mohammad Abu-Nimeh. Detecting Phishing Websites Using Machine Learning. IEEE, 2007.
5. Jain A.K. Machine Learning Techniques for Cyber Security. Springer, 2020.
6. Sahoo D. Phishing Detection Using Machine Learning Algorithms. Elsevier, 2019.
7. RFC 2827 Network Ingress Filtering.
8. OWASP Foundation Reports on Phishing Attacks.