MASHINALI O‘QITISH ALGORITMLARI YORDAMIDA TARMOQ TRAFIKIDAGI ANOMAL HOLATLARNI ANIQLASH

Yusupov Behzod Ismoil o‘g‘li

Axborot tizimlari va texnologiyalari kafedrasi assisenti

Abduraxmonova Aziza Akram qizi

Adirboyeva Shaxnoza Amir qizi

O‘zbekiston Milliy unversiteti Jizzax filliali

“Axborot tizimlari va texnologiyalari” yo‘nalishi talabalari

abduraxmonovaaziza80@gmail.com

shaxnozaadirboyeva7@gmail.com

Annotatsiya

Ushbu maqolada kompyuter tarmoqlarida anomal holatlarni aniqlashning zamonaviy usullari va gibrid yondashuvning samaradorligi tadqiq etilgan. An’anaviy xavfsizlik tizimlarining zamonaviy kiberhujumlar (DDoS, R2L, U2R) oldidagi ojizligi tahlil qilinib, mashinali o‘rganish va chuqur o‘rganish algoritmlarining afzalliklari ko‘rsatib o‘tilgan. Tadqiqot doirasida Isolation Forest, Autoencoder va Random Forest algoritmlarini birlashtirgan gibrid model taklif etilgan. Tajriba natijalari shuni ko‘rsatadiki, taklif etilgan model 97% umumiy aniqlik (Accuracy) va 0.92 F1-score ko‘rsatkichiga ega bo‘lib, tarmoq xavfsizligini ta’minlashda yuqori ishonchlilikni namoyish etadi.

Kalit so‘zlar: anomaliya aniqlash, tarmoq xavfsizligi, mashinali o‘qitish, chuqur o‘rganish, gibrid model, Isolation Forest, Autoencoder, Random Forest, DDoS hujumi, normal va anomal trafik.

Kirish

Bugungi raqamli transformatsiya davrida kompyuter tarmoqlari va internet infratuzilmalari hayotimizning ajralmas qismiga aylangan. Shu bilan birga, kiberxavfsizlik muammolari ham tobora dolzarb ahamiyat kasb etmoqda. Ayniqsa, tarmoq trafigida yuzaga keladigan anomal holatlarni aniqlash muhim vazifalardan biri hisoblanadi. Anomal

holatlar odatda tizimning normal ishlashidan og‘ishlar bo‘lib, ular ko‘pincha zararli faoliyatlar, xususan, DDoS hujumlari, ruxsatsiz kirishlar va boshqa turdagi kiberxurujlar bilan bog‘liq bo‘ladi.

An’anaviy aniqlash usullari oldindan belgilangan qoidalar asosida ishlaganligi sababli, zamonaviy va murakkab hujumlarni aniqlashda yetarli samaradorlikni ta’minlay olmaydi. Shu sababli, so‘nggi yillarda mashinali o‘qitish va chuqur o‘rganish asosidagi yondashuvlar keng qo‘llanilmoqda. Ushbu yondashuvlar katta hajmdagi ma’lumotlardan o‘rganish orqali yashirin naqshlarni aniqlash va real vaqt rejimida anomaliyalarni topish imkonini beradi.

Mazkur maqolada tarmoq trafigida anomal holatlarni aniqlash uchun gibrid model taklif etiladi. Ushbu model Isolation Forest va Autoencoder algoritmlarini birlashtirib, natijalarni Random Forest klassifikatori yordamida yakuniy qarorga keltiradi. Taklif etilgan yondashuvning asosiy maqsadi – aniqlikni oshirish, noto‘g‘ri ogohlantirishlarni kamaytirish va tizimning umumiy samaradorligini yaxshilashdan iborat.

Adabiyotlar tahlili

Anomal holatlarni aniqlash – bu tarmoqdagi odatiy faoliyatdan og‘ishgan holatlarni aniqlash jarayoni hisoblanadi. Anomal faoliyatlar asosan xakerlik hujumlari va DDoS hujumlari bilan bog‘liq bo’ladi. An’anaviy yondashuvlar esa zamonaviy tahdidlarni aniqlashga ojizlik qilmoqda. Shu sababli ham, zamonaviy tahlil metodlar: mashinali o‘rganish, chuqur o‘rganish va gibrid yondashuvlar asosiy ahamiyatga ega bo‘lmoqda. Rohan Bhagwat [1] mashinali o‘rganish usullari yordamida anomaliyalarni aniqlashda bir nechta mashinali o‘rganish algoritmlari – Logistik regressiya, SVM va Random Forest algoritmlarini o‘rgangan va tadbiq etgan. Sinovdan o‘tkazilgan modellar orasida Random Forest eng yuqori aniqlik ko‘rsatkichi va eng kam xato pozitiv (false-positive) natijalarga erishdi, bu esa uni real vaqt rejimidagi xavfsizlik ilovalari uchun ayniqsa mos qiladi. Samir Rana [2] anomaliyalarni aniqlash uchun chuqur o‘rganish va mashinani o‘rganish usullaridan foydalanish bo‘yicha va xususiyatlarni tanlash bo‘yicha tadqiqot o‘tkazgan. Algoritmlardan foydalanganda xususiyatlarni tanlash muhim omil sanaladi. Mashinali

o‘qitish yondashuvi asosida tarmoq muhitida anomal holatlarni aniqlashda Abdusadikov va boshqalar [3] gibrid model taqdim etishdi va bu model Decision Tree, SVM modellariga nisbatan aniqlik bo‘yicha va F1-score bo‘yicha yaxshiroq natijani ko‘rsatgan. Jaya Darshan va boshqalar [4] K-Means klasterlash algoritmi qo‘llanilishi o‘rganilgan. K-Means anomal trafik modellarini yetarli darajada aniqlik bilan aniqlay oladi, ayniqsa, DDoS va Probe kabi hujumlarni farqlay olishda yaxshi natija berishi mumkinligi aytib o‘tilgan. Biroq, R2L (masofadan mahalliyga) va U2R (foydalanuvchidan super-foydalanuvchiga) kabi kam uchraydigan yoki yashirin hujumlar oddiy trafikka o‘xshashligi sababli, ularni aniqlashda kamchiliklar ko‘zga tashlanadi. K-Means'ning asosiy afzalliklari uning hisoblash samaradorligi va kengayuvchanligi, bu esa uni ulkan tarmoq ma'lumotlarini real vaqtga yaqin rejimda tahlil qilishga moslashtirishi mumkinligidir. Stephanie Ness va boshqalar [5] modellarning kuchli va kuchsiz tomonlari nuqtayi nazaridan tarmoqdagi anomaliyalarni aniqlash uchun modelni to‘g‘ri tanlash zarurligini ta’kidlaydi. Isolation Forest, Naive Bayes, XGBoost, LightGBM, SVM, Random Forest va Logistik Regressiya modellarning keng turi sinovdan o‘tkazilgan.

Metodologiya

Mashinali o‘qitish bugungi raqamli davrga kelib, deyarli barcha sohalarga kirib kelgan va keng qo‘llanilib kelinmoqda. Mashinali o‘qitish (Mashine Learning) – sun’iy intellektning asosiy yo‘nalishlaridan biri va u kompyuter tizimlariga inson ishtirokisiz ma’lumotlardan o‘rganish, qaror qabul qilish imkonini beradi. Mashinali o‘qitish algoritmlari esa kompyuter dasturlarini o‘zgartirish yoki qayta dasturlashga ehtiyoj sezmasdan muayyan vazifalarni bajara oladi [7].

Isolation Forest – izolatsiya o‘rmoni anomaliyalarni aniqlash uchun ishlatiladigan foydali va samarali algoritm hisoblanadi. Ishlashi yaxshi, u xatti-harakatlarni modellashtirish o‘rniga, anomaliyalarni ularning farqlariga e’tibor qaratishi bilan ajralib turadi [8].

Random Forest – resurs talab qiluvchi: Uning aniqlash vaqti o‘rtacha bo‘lsa-da, resurs sarfi "yuqori". Bu ko‘plab qaror daraxtlari bilan ishlash jarayonida xotira va protsessor quvvatini ko‘proq sarflashi bilan izohlanadi.

SVM – og‘ir algoritm: SVM ham "Sekin" ishlaydi, ham resurs sarfi "Juda yuqori". Bu katta tarmoq tugunlarida SVMni qo‘llash operatsion jihatdan qiyin va qimmat bo‘lishini anglatadi.

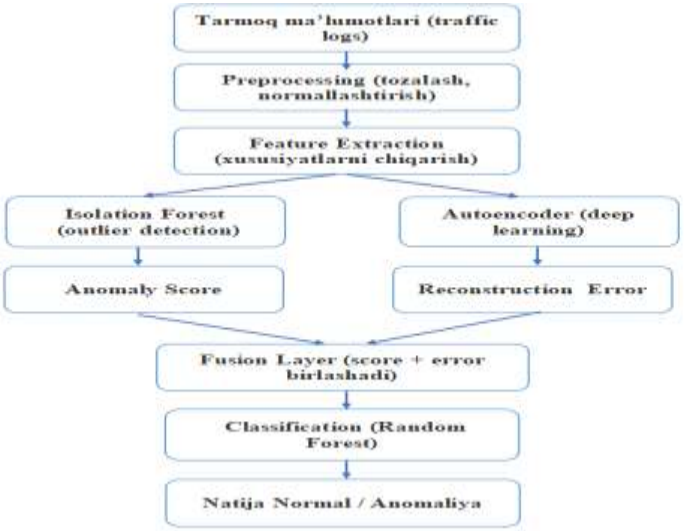
Autencoder – bu kirish ma’lumotlarini kichikroq vakillikka siqib chiqaradigan va keyin uni qayta tiklaydigan neyron tarmoqlar bo‘lib, bu modelga muhim naqshlarni o‘rganishga yordam beradi [9].

Ushbu algoritmlar tarmoqda anomal holatlarni aniqlashda quyidagi natijalarni beradi (1-jadval).

Metrika	Aniqlik	F1-score	Aniqlanish vaqti	Resurs sarfi
Autencoder	93%	0.94	O‘rtacha sekin	Yuqori
Isolation forest	90%	0.90	Tez	Past
Random forest	92%	0.91	O‘rtacha	O‘rtacha yuqori
SVM	88%	0.88	Sekin	Yuqori

1-jadval

Ma’lumotlarni integratsiyalashda va oldindan qayta ishlashda Network Traffic Anomaly Detection Dataset asosida gibrid model ma’lumotlar to‘plami shakllantirildi.



Taqdim etilayotgan model quyidagi bosqishlarni o‘z ichiga oladi:

Ma’lumotlarni yig‘ish va oldini-tayyorlash

Tarmoq trafikidan (masalan, pcap-fayllar) paket darajasidagi ma’lumotlar yig‘iladi: manzil TCP/IP, port raqamlari, paket hajmi, oqim vaqt oraliqlari, protokol turi va boshqalar.

Qo‘shimcha sifatida bir soniyadagi paketlar soni, oqim uzunligi, kirish-chiqish nisbati, retsion paketlari, katta hajmdagi paketlar soni.

Ma’lumotlarni oldindan qayta ishlash

Ma’lumotlar tozalanadi: kirishda bo‘sh qiymatlar olib tashlash, xususiyatlar normallashtirish va test- o‘qitish to‘plamlariga bo‘lish.

Xususiyatlarni tanlash va kamaytirish

Oldingi tadqiqotlarda ko‘rsatilishicha, barcha xususiyatlarni modelga kiritish aniqlikni oshirmay, balki modelni sekinlashtirishi va “overfitting” ga olib kelishi mumkin [2].

Ma’lumotlar fazosi va Vektorli ko‘rinish

Har bir tarmoq paketi (yoki oqimi) n ta xususiyatdan iborat bo‘lgan vektor ko‘rinishida ifodalanadi:

$$X_i = [x_{i1}, x_{i2}, \ldots. x_{in}]$$

Bunda:

x_{i1} – paket hajmi;

x_{i2} – vaqt oralig‘i;

x_{in} – protokol turi va boshqalar.

Normalizatsiya (Z-score Standartizatsiyasi)

Turli o‘lchov birligiga ega xususiyatlarni (masalan, paket hajmi baytlarda va vaqt millisekundlarda) bir xil miqyosga keltirish uchun quyidagi formula qo‘llaniladi

$$z = \frac{x - \mu}{\sigma}$$

Bu yerda:

μ – xususiyatning o‘rta arifmetik qiymati

σ – o‘rtacha kvadratik chetlashish.

Chegarani aniqlash va chuqur o‘qitish

Isolation Forest (Izolyatsiya o‘rmoni) algoritmi orqali odatiy ma’lumotlarga qaraganda anomaliyalar tezroq ajralib chiqadi.

Autoencoder – deep learning modeli bo‘lib, u ma’lumotlarni siqadi (encode) va yana qayta tiklaydi (decode). Bu model yashirin naqshlarni o‘rganadi va murakkab bog‘lanishlarni topadi va noliner (chiziqsiz) munosabatlarni tushunadi.

Bu ikkala modelni ishlatish yuqori aniqlik va kam xatolikni olib keladi.

Birlashtirish

Fusion Layer bu ikki xil modeldan ya’ni Isolation Forest va Autoencoderdan olingan natijalarni bitta umumiy qarorga olib kelish jarayonidir. Fusion Layer modellardan olingan natija anomaly_score va reconstruction_error ni keyingi bosqichga anomaliyani aniqlash uchun birlashtirib, klassifikatsiya funksiyasiga uzatadi.

Klassifikatsiya funksiyasi (Random Forest)

Agar N ta daraxt bo‘lsa, model daraxtlarining natijalarini umumlashtiradi. Yakuniy qaror ko‘pchilik ovoz berish (Majority Voting) prinsipi asosida qabul qilinadi.

$$Yakuniy\ bashorat = mode(T_1, T_2, \dots, T_N\ )$$

Bunda eng ko‘p uchragan qiymat olinadi.

Modelni samaradorligi quyidagi metrikalar orqali baholanadi:

1. **Aniqlik:** $P = \frac{TP}{TP+FP}$

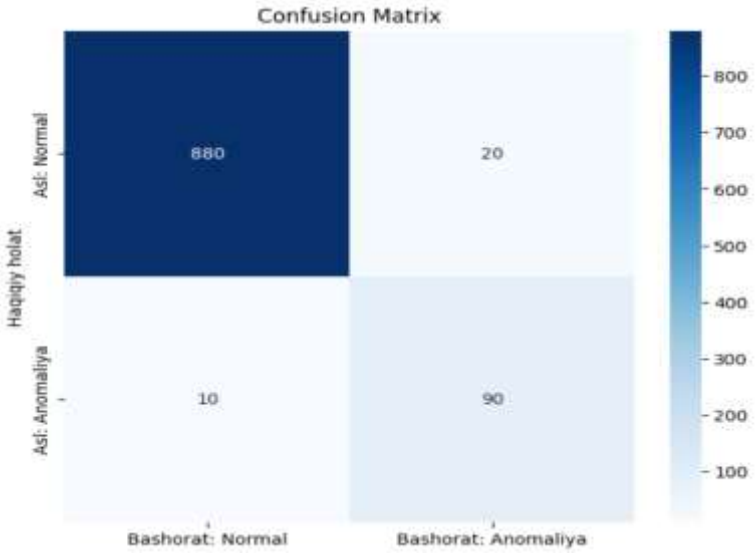
2. **To‘liqlik:** $R = \frac{TP}{TP+FN}$

3. **F1-Score:** $F1 = \frac{2*P*R}{P+R}$

TP – to‘g‘ri topilgan hujumlar, FP – xato topilgan (aslida normal) trafik, FN – o‘tkazib yuborilgan hujumlar.

Natija. Taklif etilgan gibrid model tarmoq trafigini ikki qismga normal trafik va anomal trafikka ajratadi.

Gibrid modelning Confusion Matrix (chalkashlik matrisasi). Ushbu model tarmoq hujumlarining 90% ni aniqlashga qodir. 20 ta holatda soxta signal berishi mumkin bo'lsa-da, atigi 10 ta hujumni o'tkazib yuborgani modelning kiberxavfsizlik nuqtai nazaridan ancha ishonchli ekanini ko'rsatadi (2-rasm).



2-rasm. Gibrid modelning Confusion Matrix (chalkashlik matrisasi).

Klass	Precision (Aniqlik)	Recall (qamrov)	F1-score	Support (soni)
0(Normal)	0.99	0.98	0.98	900
1(Anomaliya)	0.82	0.90	0.86	100
Accuracy (umumiy aniqlik)			0.97	1000
Macro Avg	0.90	0.94	0.92	1000
Weighted Avg	0.97	0.97	0.97	1000

2-jadval. Modelning ishlash samaradorligini ko‘rsatuvchi klassifikatsiya hisoboti.

Xulosa

Mazkur tadqiqotda tarmoq trafigida anomal holatlarni aniqlash uchun mashinali o‘qitish va chuqur o‘rganish yondashuvlariga asoslangan gibrid model taklif etildi. Model Isolation Forest, Autoencoder va Random Forest algoritmlarining kombinatsiyasi asosida ishlab chiqildi hamda ularning kuchli tomonlaridan samarali foydalanildi. Tajriba natijalari shuni ko‘rsatdiki, taklif etilgan model yuqori aniqlik (97%) va F1-score ko‘rsatkichlariga erishdi. Ayniqsa, anomal holatlarni aniqlashdagi yuqori qamrov darajasi (recall) modelning amaliy ahamiyatini oshiradi. Shuningdek, noto‘g‘ri signal berish holatlari nisbatan kam bo‘lib, bu tizimning ishonchliligini tasdiqlaydi. Gibrid yondashuvdan foydalanish natijasida an’anaviy modellar bilan solishtirganda yaxshiroq natijalarga erishildi. Bu esa bir nechta algoritmlarni birlashtirish orqali murakkab va yashirin hujumlarni aniqlash imkoniyatini kengaytiradi. Kelgusida modelni yanada takomillashtirish maqsadida real vaqt rejimida katta hajmli (Big Data) ma’lumotlar oqimida ishlash qobiliyatini oshirish ustida ish olib borish rejalashtirilgan.

Foydalanilgan adabiyotlar

[1] Rohan Bhagwat Talele “Detection Of Anomalies In Network Using Machine Learning” December 2025, Volume 13, Issue 4

[2] Samir Rana “Anomaly Detection in Network Traffic using Machine Learning and Deep Learning Techniques” Turkish Journal of Computer and Mathematics Education Vol.10 No.02 (2019), 1063-1067

[3] Abdusadikov M.A, Fozilov F.F “Tarmoq muhitida anomal holatlarni aniqlash usullari va real vaqt rejimida tarmoq trafigini tahlil qilish” Journal of Modern Educational Achievements Volume 4, 2025

[4] Jaya Darshan M, Raja Sankar A, S. Rajeswari, Dr. J.Hemalatha, D.Ramya “Network Traffic Anomaly Detection Using Machine Learning” IJCRT Volume 13, Issue 5 May 2025

[5] Stephanie Ness Vishwanath Eswarakrishnan, Harish Sridharan, Varun Shinde, Naga Venkata Prasad Janapareddy, And Vineet Dhanawat “Anomaly Detection in Network Traffic Using Advanced Machine Learning Techniques”IEEE

[7] K.K.Seitnazarov, A.T.Sultanbaeva, B.B.Aytmuratov “Mashinali o‘qitish turlari va ularning o‘ziga xos xususiyatlari” American Journal Of Education And Learning Volume-3 Issue-1 2025

[8] https://www.geeksforgeeks.org/machine-learning/what-is-isolation-forest/

[9] https://www.geeksforgeeks.org/machine-learning/auto-encoders/