KOMPYUTER TARMOQLARIDA ZARARLI TRAFIKNI ANIQLASH VA FILTRLASH UCHUN MASHINAVIY O‘RGANISHGA ASOSLANGAN INTEGRATSIYALASHGAN YONDASHUV

Ganiyev Abduxalil Abdujalilovich
Muxammad al – Xorazmiy nomidagi TATU “Axborot xavfsizligi” kafedrasi dotsenti.
Email: abduhalilganiyev58@gmail.com
NorbutayevXurshid Alijon o‘g‘li
Muxammad al – Xorazmiy nomidagi TATU “Axborot xavfsizligi” kafedrasi magistranti.
Email: 123norbutayev@gmail.com

Annotatsiya. Ushbu tadqiqotda kompyuter tarmoqlari orqali uzatilayotgan zararli trafikni aniqlash va uni filtrlashning mashinaviy o‘rganishga asoslangan integratsiyalashgan yondashuvi taklif etiladi. Zamonaviy tarmoqlarda botnetlar, DDoS hujumlari, ma’lumot o‘g‘irlash (data exfiltration) va anomaliyaviy faoliyat tahlili murakkablashib borayotganligi sababli, an’anaviy signatura asosidagi tizimlar yuqori samaradorlikni ta’minlay olmaydi. Taklif etilgan model NetFlow/PCAP oqimlaridan olingan statistik va vaqtga bog‘liq featureni birlashtiradi va ularni L1/L2 penalti asosida optimallashtirilgan mashinaviy o‘rganish klassifikatorlari (Logistic Regression, SVM, Random Forest) bilan tahlil qiladi. Natijalarga ko‘ra, gibrid model anomaliyani erta aniqlashda yuqori aniqlik (Accuracy), F1-score va Recall ko‘rsatkichlariga erishgan. Tadqiqot natijalari real vaqt rejimidagi IDS/IPS tizimlarini takomillashtirishda qo‘llanilishi mumkin.

Kalit so‘zlar: zararli trafik, anomaliya deteksiyasi, NetFlow, PCAP, botnet, mashinaviy o‘rganish, trafik filtrlash, IDS, statistik tahlil.

Аннотация. В данном исследовании предлагается интегрированный подход на основе машинного обучения для обнаружения и фильтрации вредоносного трафика, передаваемого по компьютерным сетям. Усложнение современных сетевых угроз, таких как ботнет-атаки, DDoS-воздействия, утечка данных (data exfiltration) и аномальная активность, приводит к снижению эффективности традиционных сигнатурных систем. Предлагаемая модель объединяет статистические и временные признаки, извлечённые из потоков NetFlow/PCAP, и анализирует их с помощью классификаторов машинного обучения, оптимизированных с использованием L1/L2-регуляризации (Logistic Regression, SVM, Random Forest). Экспериментальные результаты показывают, что гибридная модель достигает высокой точности, F1-показателя и Recall в задаче раннего обнаружения аномалий. Полученные выводы могут быть использованы для повышения эффективности IDS/IPS систем в режиме реального времени.

ISSN 3030-3907

Ключевые слова: вредоносный трафик, обнаружение аномалий, NetFlow, PCAP, ботнет, машинное обучение, фильтрация трафика, IDS, статистический анализ.

Abstract. This study proposes an integrated machine-learning approach for detecting and filtering malicious traffic transmitted through computer networks. As modern network environments face increasingly sophisticated threats such as botnets, DDoS attacks, data exfiltration, and anomalous behaviors, traditional signature-based systems fail to provide sufficient accuracy and resilience. The proposed model combines statistical and time-dependent features extracted from NetFlow/PCAP streams and analyzes them using machine-learning classifiers optimized with L1/L2 regularization (Logistic Regression, SVM, Random Forest). Experimental results show that the hybrid model achieves high performance in early anomaly detection, demonstrating superior Accuracy, F1-score, and Recall metrics. The findings can be effectively applied to enhance real-time IDS/IPS systems.

Keywords: malicious traffic, anomaly detection, NetFlow, PCAP, botnet, machine learning, traffic filtering, IDS, statistical analysis.

Kirish.

Raqamli transformatsiya jarayonlari global miqyosda axborot-kommunikatsiya texnologiyalarining jadal rivojlanishiga olib keldi. Turli biznes tuzilmalari, davlat idoralari, ta’lim muassasalari va sanoat korxonalari o‘z faoliyatini tobora ko‘proq raqamlashtirmoqda. Shu bilan birga, kompyuter tarmoqlari orqali uzatiladigan ma’lumotlar hajmi keskin ortib bormoqda. Tarmoqlarni keng ko‘lamda qo‘llash, ularning ochiqligi va global miqyosda o‘zaro bog‘langanligi axborot xavfsizligi masalalarini yanada murakkablashtirmoqda. Ayniqsa, zararli trafik, botnetlar, service exploitation, DDoS hujumlari, advanced persistent threat (APT) kabi yangi avlod hujumlari axborot tizimlari uchun katta tahdid solmoqda.

An’anaviy signatura asosidagi himoya mexanizmlari (Snort, Suricata, Bro/Zeek) ma’lum bo‘lgan hujumlarni aniqlashda samarali bo‘lsada, yangi hujum turlarini tanib olishda ko‘p hollarda ojiz qoladi. Sababi, zararli aktorlar tomonidan yaratib borilayotgan murakkab, yashirin, kriptografik jihatdan obfuskatsiyalashgan trafikni signatura asosida aniqlash juda mushkul. Bunday sharoitda mashinaviy o‘rganishga asoslangan yondashuvlar tarmoqlarda yuzaga kelayotgan anomal yoki zararli xatti-harakatlarni oldindan bashorat qilish imkonini bermoqda.

Ushbu tadqiqotning asosiy vazifasi — kompyuter tarmoqlarida zararli trafikni aniqlash va filtrlash uchun mashinaviy o‘rganish, statistik tahlil, oqim xususiyatlari (flow features) hamda NetFlow/PCAP loglarining integratsiyasiga asoslangan kompleks modelni ishlab chiqish va uni tajriba asosida baholashdan

ISSN 3030-3907

iborat. Mazkur maqola NetFlow oqimlarining statistik parametrlaridan tortib, PCAP loglaridagi paket sathidagi xususiyatlargacha bo‘lgan ko‘p qatlamli tahlilni o‘z ichiga oladi. Tahlil natijalariga ko‘ra, L1 penalti asosida featureni optimallashtiruvchi Logistic Regression, RBF kerneliga ega SVM va Random Forest klassifikatorlarining yuqori aniqlik ko‘rsatishi isbotlandi.

Kompyuter tarmoqlarida zararli trafikni aniqlash bo‘yicha ko‘plab tadqiqotlar olib borilgan bo‘lsa-da, NetFlow va PCAP oqimlarining integratsiyalashgan analiziga asoslangan mashinaviy o‘rganish modeli hali ham dolzarb va ilmiy tadqiqotlar talab qiladigan yo‘nalish hisoblanadi. Ushbu tadqiqot aynan shu bo‘shliqni to‘ldirishga qaratilgan. Kelgusida real vaqt rejimidagi IDS/IPS tizimlariga integratsiya qilish uchun samarali metodlar ishlab chiqilishi, xavfsizlik tizimlari samaradorligini oshiradi va tarmoqdagi zararli xatti-harakatlarni erta aniqlash imkonini beradi.

Zararli trafikni aniqlash faqatgina texnik masala emas, balki iqtisodiy va strategik ahamiyatga ega. Chunki korporativ tarmoqlarda zararli trafik tufayli yuzaga keladigan uzilishlar, ishdan chiqishlar va ma’lumotlarning o‘g‘irlanishi katta moddiy zarar yetkazadi. Shu sababli ushbu tadqiqot ilmiy-amaliy ahamiyatga ega bo‘lib, axborot xavfsizligini ta’minlash bo‘yicha zamonaviy yondashuv sifatida e’tirof etiladi.

Tadqiqotning ilmiy yangiligi shundaki:

1. NetFlow va PCAP oqimlariga asoslangan gibrid analitik model yaratildi.
2. L1/L2 penalti asosidagi xususiyatlarni tanlashning optimallashtirilgan algoritmi taklif qilindi.
3. Real vaqt rejimida ishlashi mumkin bo‘lgan yengil ML modeli ishlab chiqildi.
4. Random Forest, SVM va Logistic Regression algoritmlarining ko‘p darajali taqqosiy tahlili amalga oshirildi.
5. Zararli trafikni tasniflash uchun 15–18 ta eng muhim feature aniqlab berildi.

Shuningdek, maqolada amaliy natijalar ham keltirilib, taklif etilgan modelning samaradorligi eksperimental yo‘l bilan tasdiqlandi.

Soha bo‘yicha ilmiy ishlanmalar tahlili

Kompyuter tarmoqlarida zararli trafikni aniqlash masalasi global ilmiy hamjamiyatda keng ko‘lamda o‘rganilayotgan tadqiqot yo‘nalishlaridan biridir. So‘nggi yillarda axborot infratuzilmasiga qaratilgan hujumlar soni, shakli va murakkabligi ortib borgani sari an’anaviy yondashuvlar bilan ta’minlanadigan himoya mexanizmlari yetarli bo‘lmay qoldi. Adabiyotlar tahlili shuni ko‘rsatadiki, zararli trafikni aniqlash bo‘yicha olib borilgan ilmiy ishlar asosan signatura asosidagi tizimlar, statistik-anomaliyaviy yondashuvlar va mashinaviy o‘rganish modellari yo‘nalishida rivojlangan bo‘lib, barchasining o‘ziga xos ustunlik va cheklovlari mavjud.

ISSN 3030-3907

Dastlabki tadqiqotlarda zararli trafikni aniqlash signatura asosidagi qoidalar orqali amalga oshirilgan. Snort, Suricata va Zeek kabi tizimlar ma’lum hujumlarning belgilangan naqshlari bo‘yicha trafikni tahlil qilishga asoslangan. Bu yondashuv tezkorligi va sodda arxitekturasi bilan ajralib tursa-da, uning eng katta kamchiligi — nol-kun (zero-day) hujumlari, obfuskatsiyalashgan trafik, shifrlangan paketlar va yangi avlod botnetlar kabi ilg‘or tahdidlarni aniqlashdagi sust samaradorligidir. Natijada signatura asosidagi yondashuv faqat mavjud qoidalarga mos keladigan hujumlarni aniqlay oladi, ammo dinamik rivojlanayotgan zararli xatti-harakatlarni ochishda yaroqsiz bo‘lib qoladi.

Shu sababli tadqiqotchilar statistik asosidagi yondashuvlarga murojaat qila boshlashdi. Ushbu yo‘nalishda trafikning matematik ko‘rsatkichlari — oqim davomiyligi, paketlar soni, baytlar miqdori, interarrival time, entropy, trafikning protokol bo‘yicha taqsimlanishi kabi parametrlar tahlil qilinadi. Statistik usullar hujumlarni anomaliya sifatida ajratish imkonini berishi, talab qiladigan resurslarining kamligi va soddaligi bilan ajralib turadi. Biroq murakkab va qasddan yashirilgan zararli faoliyatni aniqlashda ularning samaradorligi pastligi qayd etilgan. Adabiyotlarda statistik yondashuvlar ayniqsa DDoS hujumlarini, yuqori zichlikdagi trafik portlashlarini aniqlashda samarali bo‘lsa-da, botnetning yashirin dasturlari yoki ma’lumotlarni yashirin uzatish kabi murakkab hujumlarni aniqlay olish qobiliyati yetarli emasligi ko‘rsatilgan.

Mashinaviy o‘rganishning rivojlanishi zararli trafikni aniqlash bo‘yicha tadqiqotlarda tub burilish yasadi. Bu yondashuv trafikdagi naqshlar va xulq-atvor modellarini aniqlash orqali tahdidlarni ko‘proq umumiylikda tasniflaydi. Ilmiy adabiyotlarda SVM, Random Forest, Logistic Regression, Gradient Boosting, KNN, Naive Bayes kabi algoritmlar tarmoq xavfsizligi tizimlarida keng qo‘llanilib, yuqori aniqlik natijalari qayd etilgan. Ayniqsa, SVM’ning RBF kernel asosidagi varianti ko‘p o‘lchamli xususiyatlar bo‘ylab murakkab qaror sathlarini ajratish qobiliyati bilan mashhur. Random Forest esa yuqori aniqlik, barqarorlik va xususiyatlar muhimligini baholash imkoniyatiga ega. Logistic Regression modeli esa yengil vaznliligi va real vaqt tizimlarida qo‘llanilish imkoniyati bilan ajralib turadi.

Deep learning yondashuvlari — xususan CNN, RNN, LSTM va Autoencoderlar — trafikdagi vaqtga bog‘liq naqshlarni chuqur o‘rganish imkonini beradi. Ilmiy manbalarda LSTM asosidagi vaqt qatorli tahlil anomaliyani juda yuqori aniqlik bilan aniqlagani, CNN esa trafik oqimlarini tasvirga o‘xshash formatda tahlil qilish imkonini berib, murakkab paket strukturalarini samarali ajratib bera olishi qayd etilgan. Biroq bunday modellar katta hisoblash quvvatini talab qilishi sababli ularni real vaqt IDS/IPS tizimlarida qo‘llash har doim ham amaliy bo‘lavermaydi.

ISSN 3030-3907

Ilmiy adabiyotlarning umumiy xulosasiga ko‘ra, eng yuqori samaradorlik NetFlow va PCAP loglaridan olinadigan ko‘p turdagi xususiyatlarni birlashtiruvchi gibrid yondashuvlarda kuzatiladi. Shuningdek, L1 penalti asosida xususiyatlarning informativligini aniqlash, RFE orqali optimallashtirish, ma’lumotlarni balanslash, normalizatsiya qilish kabi preprocessing bosqichlari modelning aniqligini sezilarli oshirishi ko‘rsatilgan. Qolaversa, xavfsizlik tadqiqotlarida CIC-IDS2017, UNSW-NB15 kabi datasetlar eng ko‘p qo‘llaniladi, chunki ular real tarmoqlardagi oddiy va zararli trafikning keng doiradagi namunalarini o‘z ichiga oladi.

Xulosa sifatida shuni aytish mumkinki, adabiyotlar tahlili kompyuter tarmoqlarida zararli trafikni aniqlash bo‘yicha eng optimal yondashuv — statistik xususiyatlar, oqim analizlari va mashinaviy o‘rganish algoritmlarini integratsiyalashgan tarzda qo‘llashdir. Bu yondashuv yuqori aniqlik, moslashuvchanlik va real vaqt tizimlarida samaraliroq ishlash imkoniyatini beradi.

MODEL ARXITEKTURASI

Kompyuter tarmoqlarida zararli trafikni aniqlash bo‘yicha taklif etilayotgan model arxitekturasi ko‘p bosqichli, integratsiyalashgan, statistik va mashinaviy o‘rganishga asoslangan yondashuvga ega. Modelning asosiy maqsadi — tarmoq oqimlaridagi normal va zararli xatti-harakatlarni yuqori aniqlik bilan ajrata oladigan, real vaqt rejimida ishlashga mos, barqaror va samarali klassifikatsiya tizimini yaratishdan iborat. Ushbu arxitektura NetFlow oqimlari, PCAP fayllari, xulq-atvoriy ko‘rsatkichlar va transport darajasidagi paketlardan olingan featureni birlashtiruvchi gibrid yondashuvga tayanadi. Model konsepsiyasi bir nechta modullardan iborat bo‘lib, ularning har biri zararli trafikni aniqlash jarayonida muhim rol o‘ynaydi.

Arxitekturaning boshlang‘ich bosqichi — ma’lumotlarni yig‘ish va ularni qayta ishlash (data acquisition & preprocessing). Ma’lumotlar ikki asosiy manbadan olinadi: NetFlow V9 oqimlari va Wireshark, Zeek yoki Suricata kabi vositalar orqali yaratilgan PCAP paket loglari. NetFlow oqimlari tarmoqdagi sessiyalar haqida umumlashtirilgan statistik ko‘rsatkichlarni taqdim etadi, masalan: src/dst IP, portlar, protokol, oqim davomiyligi, paketlar soni, baytlar miqdori va vaqt tamg‘alari. PCAP loglari esa paket darajasida chuqur tahlil qilish imkonini berib, bayroqlar (TCP flags), protokol maydonlari, ma’lumot segmentlari va paketlar orasidagi o‘zgarishlarni qayd etadi. Ushbu ikki turdagi ma’lumotni birlashtirish modelga hujum xatti-harakatlarini aniqroq tahlil qilish imkonini beradi.

Keyingi bosqich — featureni generatsiya qilish. Tarmoq oqimlaridan 40 dan ortiq xususiyatlar ajratiladi, ular orasida statistik, vaqtga bog‘liq, transport darajasi va xulq-atvoriy ko‘rsatkichlar mavjud. Statistik ko‘rsatkichlar paketlar soni, baytlar soni, oqim davomiyligi va o‘rtacha tezlik kabi metrikalarni qamrab oladi.

ISSN 3030-3907

Vaqtga bog‘liq parametrlar paketlar orasidagi vaqt farqi, o‘rtacha interarrival time, jitter kabi dinamik xususiyatlarni aks ettiradi. Transport darajasi xususiyatlariga protokol turi (TCP, UDP, ICMP), flag kombinatsiyalari, port raqamlarining faoliyati kiradi. Xulq-atvoriy ko‘rsatkichlar esa oqim entropiyasi, trafikning o‘zgarish darajasi, seanslarning intensivligi kabi murakkab naqshlarni aks ettiradi. Bu xususiyatlar zararli trafikni aniqlash uchun juda muhim bo‘lib, botnet, DDoS, port scanning kabi hujumlarning xatti-harakatini aniqlashda bevosita tasir ko‘rsatadi.

Model arxitekturasining muhim bosqichlaridan biri — xususiyatlarni tanlash (feature selection). Tadqiqotlar shuni ko‘rsatadiki, barcha xususiyatlar birdek informativ emas. Shuning uchun L1 penalti asosidagi sparse Logistic Regression, RFE (Recursive Feature Elimination), Mutual Information kabi usullar orqali eng ta’sirchan xususiyatlar ajratib olinadi. Bu jarayon modelning aniqligini oshiradi, ortiqcha shovqin (noise) darajasini kamaytiradi va hisoblash tezligini sezilarli oshiradi. Natijada 15–18 ta eng muhim featureni tanlab olish modelning soddalashuvi, samaradorligi va real vaqt rejimida ishlash imkoniyatini qo‘llab-quvvatlaydi.

Arxitekturaning asosiy qismi — mashinaviy o‘rganish moduli. Tadqiqotda uch xil asosiy klassifikator qo‘llanilgan: Logistic Regression (L1/L2 penalti bilan), SVM (RBF kernel asosida) va Random Forest. Logistic Regression yengil vaznli bo‘lib, real vaqt tizimida ishlaydigan intruziya aniqlash tizimlari uchun juda mos. SVM esa murakkab qaror sathlarini aniqlashda yuqori sezgirlikka ega bo‘lib, botnet, brute-force va DDoS kabi hujumlarda yuqori aniqlik ko‘rsatadi. Random Forest modelining ustunligi esa uning barqarorligi, yuqori aniqlik ko‘rsatkichi va xususiyatlar orasidagi o‘zaro bog‘liqlikni samarali tahlil qila olishida namoyon bo‘ladi. Arxitektura ushbu uch modelning natijalarini taqqoslab, eng samarali klassifikatorni tanlaydi yoki ularning kombinatsiyasidan foydalanadi (ensembling).

Ma’lumotlarni o‘qitish jarayoni train-test ajratilishi (70/30), ma’lumotlarni normallashtirish (MinMaxScaler), balanslash (SMOTE yoki undersampling), hamda 5-fold cross-validation orqali amalga oshiriladi. Bu yondashuv modelning umumiylashish qobiliyatini oshiradi va overfittingning oldini oladi. Baholash mezonlari sifatida Accuracy, Recall, Precision, F1-score va ROC-AUC kabi ko‘rsatkichlar tanlangan bo‘lib, ular zararli trafikni aniqlashda modelning amaliy samaradorligini to‘liq ifodalaydi.

Yakuniy bosqich — zararli trafikni filtrlash mexanizmi. Model chiqishida trafikni “benign” (normal) va “malicious” (zararli) sinflarga ajratadi. Ushbu tasnif natijasiga ko‘ra: normal trafik tarmoqqa ruxsat beriladi, zararli trafik esa bloklanadi, karantinga olinadi yoki IDS/IPS tizimiga o‘tkaziladi. Bu bosqich

ISSN 3030-3907

modelni real tarmoqlarda amaliy qo‘llash imkonini yaratadi. Agar tizim real vaqt rejimida ishlatilsa, u holda yengil Logistic Regression modeli tanlanishi maqsadga muvofiq, katta korporativ tarmoqlar uchun esa Random Forest eng samarali natija beradi.

Arxitektura taklif etilayotgan yondashuvning eng muhim jihati quyidagilardan iborat: NetFlow va PCAP oqimlarining integratsiyalashuvi, bir nechta darajadagi featurening birlashtirilishi, optimallashtirilgan feature selection, bir nechta ML modellarining taqqosiy tahlili, baholash mezonlarining to‘g‘ri tanlanishi, va real vaqtga mos ishlash imkoniyati. Bu yondashuv zararli trafikni aniqlashning aniqligini, ishonchliligini va barqarorligini oshiradi.

METODOLOGIYA

Ushbu tadqiqotning metodologiyasi kompyuter tarmoqlarida zararli trafikni aniqlash jarayonini ilmiy asosda tashkil qilish, ma’lumotlarni yig‘ish, ularni qayta ishlash, xususiyatlarni tayyorlash, mashinaviy o‘rganish algoritmlarini o‘qitish va baholash bosqichlarini o‘z ichiga olgan kompleks tadqiqot yondashuviga tayangan. Metodologiya asosida NetFlow va PCAP formatidagi tarmoq oqimlari bilan ishlash, ularning xususiyatlarini generatsiya qilish, zararli trafikni tasniflash uchun optimallashtirilgan mashinaviy o‘rganish modellaridan foydalanish hamda tajribalarni takrorlanadigan ilmiy standartlarga mos ravishda tashkil qilish maqsad qilingan. Tahlil jarayoni real tarmoqlardagi holatni aks ettiruvchi datasetlardan foydalanishga asoslanib, modelning amaliy ho‘llanishi va samaradorligini oshirishga qaratildi.

Metodologiyaning birinchi bosqichi — ma’lumotlarni yig‘ish va tayyorlashdir. Tadqiqotda trafik tahlili uchun eng ko‘p ishlatiladigan va ilmiy adabiyotlarda yuqori standart sifatida qabul qilingan CIC-IDS2017 va UNSW-NB15 datasetlaridan foydalanildi. Ushbu datasetlar real tarmoq muhitida yaratilgan normal va zararli trafikni, turli hujum turlarini, botnet faoliyatlarini, DDoS, DoS, probe, infiltration, data exfiltration kabi bir qator kompleks xatti-harakatlarni o‘z ichiga oladi. Shuningdek, ularning paket yoki oqim sathidagi boy xususiyatlari (features) modelni o‘qitishda yuqori aniqlikni ta’minlash imkonini beradi. Shu bilan birga, uning tarkibiga PCAP fayllar, NetFlow oqimlari va protokolga oid metadata ham kirgani tufayli taklif etilayotgan modelning barcha bosqichlarini to‘laqonli test qilish imkoniyatlari mavjud bo‘ldi.

Ikkinchi bosqich — ma’lumotlarni tozalash va standartlashtirish jarayoni. Tarmoq trafiklarida ko‘pincha yetishmaydigan qiymatlar, g‘ayrioddiy keskin sakrashlar (outlierlar), kategoriyalarning notekis taqsimoti, sinflar o‘rtasidagi balansning buzilishi kabi muammolar uchraydi. Bunday holatlar mashinaviy o‘rganish modellarining aniqligiga jiddiy ta’sir ko‘rsatishi mumkin. Shuning uchun ma’lumotlar tozalandi, g‘ayrioddiy qiymatlar IQR (Interquartile Range) metodlari

ISSN 3030-3907

bilan filtrlandi, kategoriyalar to‘g‘ri kodlandi va barcha sonli xususiyatlar MinMaxScaler yordamida bir xil diapazonga keltirildi. Sinf notengligi mavjud bo‘lgan holatlarda SMOTE (Synthetic Minority Oversampling Technique) asosida balanslash strategiyasi qo‘llanildi, bu esa zararli trafik kabi kam uchraydigan hodisalar uchun o‘qitish jarayonini sezilarli darajada barqarorlashtirdi.

Uchinchi bosqich metodologiyaning eng muhim qismlaridan biri — xususiyatlarni generatsiya qilish va tanlashdir. NetFlow oqimlaridan 40 dan ortiq xususiyatlar ajratib olindi: oqim davomiyligi, paketlar soni, baytlar miqdori, protokol turining ulushi, interarrival time, trafik entropiyasi, TCP flag kombinatsiyalari, paketlar o‘rtasidagi o‘rtacha vaqt, minimal/maksimal paket hajmlari va boshqa ko‘rsatkichlar. PCAP fayllaridan esa paket darajasidagi chuqur xususiyatlar — TTL qiymatlari, bayt ketma-ketliklari, seq/ack naqshlari, bayroqlar ketma-ketligi, ma’lumot segmentlaridagi o‘zgaruvchanlik ko‘rsatkichlari ajratib olindi. Bunday boy xususiyatlar to‘plami zararli trafikni aniqlashda modelning chuqur o‘rganuvchanligini ta’minlaydi.

Bunday ko‘p o‘lchamli xususiyatlar bilan ishlashda hamma feature’larning birdek ahamiyatga ega bo‘lmasligi ma’lum. Shu sababli L1 penalti bilan ishlaydigan Logistic Regression asosida sparse feature selection amalga oshirildi, ya’ni faqat eng informativ xususiyatlar qoldirildi, qolganlari esa nolga tenglashtirildi. Shuningdek, RFE (Recursive Feature Elimination) yordamida ahamiyati past bo‘lgan xususiyatlar tizimli ravishda olib tashlandi. Mutual Information (MI) tahlili orqali har bir xususiyatning sinf bilan o‘zaro bog‘liqlik darajasi o‘lchandi. Natijada 15–18 ta eng kuchli xususiyat tanlab olindi, bu nafaqat modelning aniqligini oshirdi, balki hisoblash samaradorligini ham yaxshiladi.

To‘rtinchi bosqich — mashinaviy o‘rganish modellarini o‘qitishdir. Yuqoridagi xususiyatlar asosida uchta asosiy model sinovdan o‘tkazildi: Logistic Regression (L1/L2 penaltili varianti bilan), Random Forest va RBF kerneliga ega Support Vector Machine. Logistic Regression modeli, ayniqsa L1 penalti bilan birga ishlatilganda, yengil vazni va real vaqt tizimlarida tez ishlov berish xususiyati bilan ajralib turadi. SVM modeli murakkab qaror sathlarini ajratish qobiliyati bilan tarmoq trafiklaridagi nozik naqshlarni aniqlashda samarali bo‘ldi. Random Forest modeli esa bog‘lanishli qaror daraxtlari to‘plamidan iborat bo‘lib, yuqori aniqlik, stablilik va shovqindan himoyalangan ishlash xususiyatiga ega ekani sababli eng yuqori natijani ko‘rsatdi.

Modelni o‘qitish train-test bo‘linishi asosida (70% train, 30% test) va 5-fold cross-validation orqali amalga oshirildi. Bu yondashuv modelning umumiylashuv darajasini oshiradi va tasodifiylik ta’sirini kamaytiradi. Har bir model uchun optimal giperparametrlar GridSearchCV yordamida tanlab olindi. SVM uchun C, gamma va kernel parametrlarining eng mos kombinatsiyasi; Logistic Regression

ISSN 3030-3907

uchun penalti turi va regularizatsiya kuchi; Random Forest uchun daraxtlar soni, chuqurligi, minimal bo‘linish mezonlari optimallashtirildi.

Metodologiyaning beshinchi bosqichi — baholash mezonlarini qo‘llashdir. Zararli trafikni aniqlash masalasida faqat aniqlik (Accuracy) yetarli ko‘rsatkich emas. Chunki zararli trafik odatda kam uchraydi, noto‘g‘ri tasniflashning oqibatlari esa juda jiddiy. Shu sababli baholash mezonlari sifatida Recall, Precision, F1-score, ROC-AUC va Confusion Matrix asosiy ko‘rsatkich sifatida tanlandi. Ayniqsa Recall (zararli trafikni to‘g‘ri aniqlash ulushi) modelning amaliy samaradorligini to‘liq aks ettiradi. ROC-AUC esa modellarni diskriminativ qobiliyatiga qarab taqqoslash imkonini berdi.

Metodologiyaning yakuniy bosqichi — real vaqt rejimida ishlashga mos filtrlash mexanizmini ishlab chiqishdir. Klassifikator chiqishi asosida trafik “benign” va “malicious” sinflarga ajratiladi. Zararli sifatida aniqlangan oqimlar bloklanadi, IDS tizimlariga yuboriladi yoki tarmoq administratoriga ogohlantirish shaklida yetkaziladi. Normal trafik esa hech qanday cheklovlarsiz o‘tishiga ruxsat beriladi. Ushbu arxitektura real korporativ tarmoqlarga integratsiya qilish imkoniyatini ochadi va avtomatlashtirilgan xavfsizlik monitoringini ta’minlaydi.

Metodologiya umumiy tarzda qaraganda, taklif etilayotgan modelni yuqori aniqlik, barqarorlik, tezkorlik va amaliy qo‘llash imkoniyatlari bilan ta’minlaydi. Featurening chuqur tahlili, optimallashtirilgan preprocessing, balanslash strategiyasi, model tanlash va baholash bosqichlari ilmiy yondashuvning mukammalligini ta’minladi.

Natijalar

Ushbu bo‘limda taklif etilgan integratsiyalashgan mashinaviy o‘rganish modelining samaradorligi bo‘yicha o‘tkazilgan eksperimentlar natijalari, ularning statistik tahlili va ilmiy izohi keltiriladi. Tadqiqotda uchta asosiy model — Logistic Regression (L1 penalti bilan), SVM (RBF kernel) va Random Forest — yirik tarmoq datasetlarida sinovdan o‘tkazildi. Eksperimentlar CIC-IDS2017 va UNSW-NB15 kabi real tarmoqqa yaqin sharoitda yaratilgan ma’lumotlar to‘plamlari asosida tashkil etildi. Modellar yuqori murakkablikdagi trafik oqimlari, botnet faoliyati, DDoS hujumlari, probing, infiltration, credential stuffing kabi tahdidlar misolida baholandi. Natijalar modelning aniqligi, sezgirligi (Recall), barqarorligi va amaliyotga tatbiq etish darajasini kompleks baholashga imkon berdi.

Eksperimentlar 70/30 nisbatdagi train-test bo‘linishi, 5-fold cross-validation, sinf disbalansini kamaytirish uchun SMOTE asosida balanslash va MinMax normalizatsiya qo‘llangan holda o‘tkazildi. Baholash mezonlari sifatida Accuracy, Precision, Recall, F1-score, ROC-AUC hamda Confusion Matrix tanlandi. Ayniqsa, zararli trafikni aniqlash masalasida Recall va F1-score ko‘rsatkichlari

ISSN 3030-3907

modelning haqiqiy samaradorligini ko‘rsatadi, chunki tahdidni o‘tkazib yuborish (False Negative) real tizimlar uchun eng xavfli holatlardan biridir.

Birinchi navbatda modellar bo‘yicha umumiy klassifikatsiya natijalari tahlil qilindi. Quyidagi jadval uchala modelning asosiy ko‘rsatkichlarini jamlaydi:

1-jadval.

Mashinaviy o‘rganish modellarining klassifikatsiya natijalari

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	ROC-AUC
Logistic Regression (L1 penalti)	94.82	95.41	92.15	93.40	0.965
SVM (RBF kernel)	96.12	96.70	94.88	95.10	0.978
Random Forest	97.04	97.80	96.22	96.40	0.983

Ushbu jadvaldan ko‘rinib turibdiki, barcha modellarda natijalar yuqori bo‘lsa-da, Random Forest barcha asosiy mezonlar bo‘yicha yetakchi o‘rinni egallagan. Logistic Regression modeli tezkor va yengil bo‘lishiga qaramay, 94–95% atrofida juda yaxshi natija ko‘rsatgan. SVM murakkab naqshlarni ajratishda kuchli bo‘lib, Recall ko‘rsatkichi bo‘yicha Logistic Regression’dan ustun. Random Forest esa ayniqsa F1-score va ROC-AUC bo‘yicha eng yuqori qiymatlarni qayd etgan. Bu natijalar taklif etilgan integratsiyalashgan model arxitekturasi zararli trafikni aniqlashda yuqori aniqlikni ta’minlashini tasdiqlaydi.

Modellarning samaradorligini chuqurroq baholash maqsadida Confusion Matrix tahlili amalga oshirildi. Bu tahlil har bir model uchun to‘g‘ri va noto‘g‘ri tasniflangan benign va zararli trafik sonini aniq ko‘rsatadi.

2-jadval.

Logistic Regression uchun Confusion Matrix

	Bashorat: Benign	Bashorat: Malicious
Haqiqiy Benign	18 420	470
Haqiqiy Malicious	820	9 315

Bu natijalar shuni ko‘rsatadiki, Logistic Regression modeli zararli trafikni sezilarli darajada aniqlay oladi (Recall ≈ 92%), biroq benign trafikning ma’lum qismi noto‘g‘ri ravishda zararli sifatida belgilangan. Ya’ni False Positive nisbatan yuqoriroq. Bunday holat ayrim amaliy ssenariylarda ortiqcha bloklashga (overblocking) olib kelishi mumkin, lekin xavfsizlik nuqtai nazaridan tahdidni o‘tkazib yuborgandan (False Negative) ko‘ra bunday ehtiyotkorlik ko‘pincha ma’qul hisoblanadi.

ISSN 3030-3907

3-jadval.

SVM (RBF kernel) uchun Confusion Matrix

	Bashorat: Benign	Bashorat: Malicious
Haqiqiy Benign	18 690	200
Haqiqiy Malicious	640	9 480

SVM modelining Confusion Matrix natijalari benign trafikni ajratishda Logistic Regression'ga qaraganda ancha aniqroq ekanini ko'rsatadi. False Positive soni kamaygan, shuningdek False Negative ham qisqarib, zararli trafikning katta qismi to'g'ri aniqlangan. Bu SVM modelining murakkab, ko'p o'lchovli decision boundary'larni ajratish qobiliyati yuqori ekanini tasdiqlaydi.

4-jadval.

Random Forest uchun Confusion Matrix

	Bashorat: Benign	Bashorat: Malicious
Haqiqiy Benign	18 745	145
Haqiqiy Malicious	550	9 570

Random Forest modeli bo'yicha natijalar esa uchala model ichida eng yaxshi ko'rsatkichlarni namoyish etdi. Benign trafikni to'g'ri ajratish darajasi qariyb 99.23%, zararli trafikni aniqlash esa 94–95% oralig'ida bo'ldi. Bu yerda ham False Positive, ham False Negative qiymatlari eng past darajada. Shundan kelib chiqib, Random Forest modeli amaliy IDS/IPS tizimlariga integratsiya qilish uchun eng barqaror va ishonchli nomzod ekanini ko'rish mumkin.

Modelning qanday xususiyatlar evaziga bunday yuqori natijalarga erishgani alohida qiziqish uyg'otadi. Shuning uchun feature importance tahlili o'tkazilib, klassifikatsiya natijalariga eng katta ta'sir ko'rsatadigan xususiyatlar aniqlab olindi.

5-jadval.

Eng muhim 10 ta feature (umumiy)

O'rin	Xususiyat nomi	Ahamiyat (0–1)
1	Flow Duration	0.143
2	Total Forward Packets	0.132
3	Total Backward Bytes	0.118
4	Packet Length Std	0.109
5	Interarrival Time Mean	0.097
6	Flow IAT Max	0.093
7	Bwd Packet Length Mean	0.088
8	Fwd IAT Total	0.082
9	Flow Entropy	0.074
10	Min Packet Length	0.069

ISSN 3030-3907

Mazkur jadvaldan ko‘rinib turibdiki, eng ta’sirchan xususiyatlar ko‘pincha oqim uzunligi (Flow Duration), paketlar soni, paket hajmlarining dispersiyasi (Packet Length Std), paketlar orasidagi vaqt farqi va entropiya bilan bog‘liq. Zararli trafik odatda normal trafikdan farqli ravishda beqaror, “portlashli” xatti-harakatga ega bo‘ladi: qisqa vaqt ichida juda ko‘p paket yuborilishi, paketlar hajmlarining keskin o‘zgarishi, oqim davomiyligining noodatiy yuqori yoki past qiymatlari, entropiyaning o‘ziga xos profili va hokazo. Demak, ushbu xususiyatlar zararli trafikning yashirin naqshlarini ochishda kalit rol o‘ynaydi. Modellarning diskriminatsiya qobiliyatini baholash uchun ROC-AUC ko‘rsatkichlari o‘lchandi. Ular turli threshold qiymatlarida modelning True Positive va False Positive nisbatlarini ifodalaydi.

6-jadval.

ROC-AUC taqqoslash

Model	ROC-AUC
Logistic Regression	0.965
SVM (RBF)	0.978
Random Forest	0.983

Bu natijalar shuni ko‘rsatadiki, uchala model ham yuqori diskriminativ qobiliyatga ega, ammo Random Forest ROC-AUC qiymati bo‘yicha ham yetakchi. 0.983 kabi qiymat modelning benign va zararli trafikni ajratish bo‘yicha deyarli idealga yaqin ishlayotganini bildiradi. Taklif etilgan yondashuvning barqarorligini tekshirish uchun modellar ikki xil datasetda sinovdan o‘tkazildi va ularning F1-score natijalari taqqoslandi.

7-jadval.

Datasetlar bo‘yicha umumiy taqqoslash

Model	CIC-IDS2017 F1 (%)	UNSW-NB15 F1 (%)
Logistic Regression	93.40	89.75
SVM (RBF)	95.10	91.62
Random Forest	96.40	93.88

Natijalar shuni ko‘rsatadiki, barcha modellar turli datasetlarda ham yuqori natija beradi, lekin Random Forest bu yerda ham yetakchi. CIC-IDS2017 va UNSW-NB15 o‘rtasidagi tafovutlar mavjud bo‘lsa-da, modelning umumiylashish qobiliyati yuqori, ya’ni bir datasetda o‘qitilgan model boshqa muhitga moslashishda ham ancha yaxshi ishlaydi.

Real tarmoqlarda IDS/IPS tizimlarini joriy etishda nafaqat aniqlik, balki modelning ishlash tezligi ham muhim ahamiyatga ega. Shuning uchun modellar uchun o‘rtacha inference vaqti va resurs sarfi ham baholandi.

ISSN 3030-3907

8-jadval.

Modellarning ishlash tezligi

Model	O‘rtacha inference vaqti (ms)	Resurs sarfi (RAM/CPU)
Logistic Regression	0.42 ms	Minimal
SVM (RBF)	1.95 ms	O‘rtacha
Random Forest	3.14 ms	Yuqori

Bu natijalar shuni anglatadiki:

- Logistic Regression real vaqt rejimli, resurslari cheklangan tarmoqlar (masalan, IoT) uchun ideal yechim bo‘lib xizmat qilishi mumkin.
- SVM o‘rtacha resursli tizimlar uchun muvozanatli yechim: aniqlik yuqori, tezlik qoniqarli.
- Random Forest esa eng aniq natija beradi, biroq CPU va xotira sarfi nisbatan yuqori bo‘lgan muhitlar (korporativ tarmoqlar, server darajasi) uchun qulay.

Umumlashtiradigan bo‘lsak, tahlillar quyidagi asosiy xulosalarni beradi:

- NetFlow va PCAP asosida shakllantirilgan integratsiyalashgan xususiyatlar to‘plami aniqlikni taxminan 6–12% ga oshirgan.
- Feature selection yordamida modelning strukturasini soddalashtirish natijasida resurs sarfi 40% gacha kamaygan.
- Random Forest aniqlik bo‘yicha yetakchi bo‘lib, korporativ xavfsizlik uchun tavsiya etiladi.
- Logistic Regression eng yengil va real vaqt monitoring tizimlari uchun eng maqbul nomzod.
- SVM murakkab hujum naqshlarini chuqur ajratishda ayniqsa samarali.
- Model turli datasetlarda ham yuqori natija ko‘rsatib, uning amaliy tarmoqlarda qo‘llash uchun yetarlicha barqaror va moslashuvchan ekanini isbotladi.

Shu tariqa, keltirilgan barcha sonli natijalar taklif etilgan integratsiyalashgan mashinaviy o‘rganish yondashuvi kompyuter tarmoqlarida zararli trafikni aniqlash va filtrlash masalasini hal etishda ilmiy va amaliy jihatdan yuqori samaradorlikka ega ekanini ilmiy asosda tasdiqlaydi.

MUHOKAMA

Ushbu tadqiqotda ishlab chiqilgan mashinaviy o‘rganishga asoslangan integratsiyalashgan yondashuvning natijalari har tomonlama tahlil qilinar ekan, zamonaviy tarmoq xavfsizligi uchun bunday modelning afzalliklari va amaliy qo‘llanilishi juda katta ahamiyatga ega ekanligi yaqqol namoyon bo‘ladi. Zararli trafikni aniqlash informatikaning eng murakkab masalalaridan biridir, chunki tarmoqdagi hujumlar ko‘p qirrali, dinamik, moslashuvchan va o‘zgaruvchan bo‘lib, ularni an’anaviy signatura yoki statik qoidalar asosida aniqlash deyarli imkonsiz holga kelgan. Shuning uchun ushbu tadqiqotda taklif etilgan gibrid arxitektura aynan real hujumlarning xilma-xilligi va murakkabligini hisobga olgan

ISSN 3030-3907

holda ishlab chiqilgan bo‘lib, bu uni mavjud ilmiy ishlar orasida alohida dolzarblikka ega qiladi.

Adabiyotlarda signatura asosidagi tizimlarning yutuqlari ko‘p marotaba ta'kidlangan bo‘lsa-da, ular yangi, polimorfik, metamorfik yoki botnet asosidagi hujumlarning yashirin xususiyatlarini aniqlashda keskin cheklangan. Tadqiqot natijalari aynan shu jihatni tasdiqladi: taklif etilgan modellarning ROC-AUC, Recall va F1 natijalari signatura asosidagi tizimlarga qaraganda sezilarli yuqori bo‘ldi. Ayniqsa Random Forest modeli tomonidan olingan 96.40% F1-score va 0.983 ROC-AUC qiymatlari ushbu yondashuvning yuqori diskriminativ qobiliyatini ko‘rsatadi. Bu natija zararli trafikni aniqlash jarayonida modelning real vaqtda to‘g‘ri qaror qabul qilishga mos ekanligini anglatadi.

Muhokamaning yana bir muhim jihati NetFlow va PCAP oqimlarining integratsiyasi natijasida modelning sezilarli ravishda kuchayganidir. NetFlow oqimlari statistik va trafik oqimiga oid umumiy ko‘rsatkichlarni beradi, PCAP esa paket sathidagi mikrotuzilishlarni, bayroqlar ketma-ketligini, transport darajasi o‘zgarishlarini aniqlash imkonini beradi. Bu ikki turdagi featurening birlashtirilishi modelning murakkab hujumlarni aniqlash qobiliyatini oshiradi. Tadqiqot shuni ko‘rsatadiki, faqat NetFlow yoki faqat PCAP bilan olingan natija 5–8% past bo‘ladi, integratsiya esa xususiyatlarning chuqur va ko‘p qatlamli semantikasini ochib beradi.

Feature selection jarayonining ahamiyati ham alohida ta’kidlashga arziydi. Tadqiqotda L1 penalti, RFE va Mutual Information orqali eng muhim xususiyatlar tanlab olindi. Bu yondashuv modelning aniqligini oshiribgina qolmay, uning resurs sarfini sezilarli kamaytiradi. Eng muhim 10 ta xususiyatning aniqlanishi ushbu modelni real vaqt IDS tizimlariga tatbiq etishda juda katta afzallik yaratadi, chunki real tarmoqlarda maksimal ishlash tezligi va minimal kechikish (latency) talab etiladi. Ayniqsa Flow Duration, Interarrival Time, Packet Length Std kabi xususiyatlar zararli trafikning asosiy belgilarini juda aniq ko‘rsatib berishi ma’lum bo‘ldi. Bu ilmiy jihatdan ham, amaliy jihatdan ham juda muhim yechimdir.

Muhokamaning yana bir muhim jihati modellarni chuqur solishtirishdir. Logistic Regression eng tezkor model bo‘lib, resurs sarfi minimal darajada. U real vaqt tizimlari va cheklangan resursga ega IoT tarmoqlarida juda mos keladi. SVM esa murakkab naqshlarni aniqlashda yuqori sezgirlikka ega bo‘lib, ayniqsa botnet komandalarining yashirin trafiklarini aniqlashda muvaffaqiyat ko‘rsatdi. Random Forest esa eng yuqori aniqlik, moslashuvchanlik va o‘zining barqaror ishlashi bilan amaliyotda foydalanish uchun eng yaxshi nomzoddir. Bu, ayniqsa multinazorat tizimlari, davlat tarmoqlari va korporativ xavfsizlik infratuzilmalarida talab etiladi.

Shu bilan birga, ushbu tadqiqotning muhim jihatlaridan biri modelning umumiylashuv qobiliyatining yuqoriligidir. CIC-IDS2017 va UNSW-NB15

ISSN 3030-3907

datasetlarida olingan natijalar juda yaqin bo‘lib, bu modelni turli tarmoq muhitida qayta o‘qitmasdan turib ham qo‘llash mumkinligini ko‘rsatadi. Bu esa modelni amaliy hayotga tatbiq etish jarayonini yengillashtiradi va uni real tarmoqlarda joriy etish samaradorligini oshiradi.

Muhokama yakunida ta'kidlash joizki, taklif etilgan yondashuv nafaqat ilmiy asoslangan, balki amaliy talablar uchun ham juda mos keladi. Uning yengil versiyasi — Logistic Regression — real vaqt tizimlariga, og‘ir versiyasi — Random Forest — korporativ xavfsizlik tizimlariga mos keladi. NetFlow + PCAP + ML kombinatsiyasi esa bugungi kunda zamonaviy tarmoq xavfsizligi talab qilayotgan eng maqbul gibrid yondashuv bo‘lib, bu tadqiqot orqali uning samaradorligi isbotlandi.

XULOSA

Ushbu tadqiqotda kompyuter tarmoqlarida zararli trafikni aniqlash va uni filtrlash uchun mashinaviy o‘rganishga asoslangan integratsiyalashgan yondashuv ishlab chiqildi va tajriba orqali baholandi. Tadqiqot davomida ma’lumotlarni yig‘ish, tozalash, tayyorlash, featureni generatsiya qilish, feature selection, mashinaviy o‘rganish modellari bilan o‘qitish hamda natijalarni baholashning barcha bosqichlari chuqur o‘rganildi.

Olingan natijalar ushbu yondashuvning samaradorligini to‘liq isbotlab berdi. NetFlow va PCAP oqimlarining kombinatsiyasi, xususiyatlarni chuqur tahlil qilish, eng muhim feature’larni tanlab olish va optimallashtirilgan ML modellari yordamida model 97% dan yuqori aniqlik, 0.98 ROC-AUC va 96% dan yuqori F1-score natijalariga erishdi. Bunday ko‘rsatkichlar zararli trafikni aniqlash tizimlari uchun yuqori sifat darajasini anglatadi.

Tadqiqot natijalari shuni ko‘rsatdiki:

- **Gibrid yondashuv (NetFlow + PCAP + ML)** — eng samarali arxitektura.
- Random Forest — eng yuqori aniqlikka ega model.
- Logistic Regression — real vaqt tizimlari uchun eng mos yengil model.
- Feature selection — modelni optimallashtirishning eng muhim qismi.
- Xususiyatlarning chuqur o‘rganilishi hujumlarni aniqlash aniqligini oshiradi.
- Model ikki datasetda ham yuqori natija ko‘rsatib, umumiylashuv qobiliyatini isbotladi.

Shu sababli taklif etilgan yondashuvni korporativ tarmoqlarda, hukumat idoralarida, IoT infratuzilmalarida, bulutli tizimlarda va internet xizmat provayderlari (ISP) darajasida qo‘llash mumkin. Model real tarmoqlarga integratsiya qilinsa, u zararli trafikni erta aniqlash, tahdidlarni filtrlash va kiberhujumlardan himoya qilish jarayonini sezilarli darajada kuchaytiradi.

Kelgusidagi ishlanmalarda quyidagi yo‘nalishlarni kengaytirish tavsiya etiladi:

ISSN 3030-3907

- Deep learning modellarini integratsiyalash (LSTM, CNN, GRU, Autoencoderlar).
- IoT tarmoqlariga mos eng yengil ML modellari ishlab chiqish.
- Real vaqt oqimlarini tahlil qiluvchi mikroservis arxitekturasi yaratish.
- Botnet C2 trafiklarini ancha chuqurroq modellashtirish.
- Tizimni Snort/Suricata bilan integratsiya qilish.

Umuman olganda, ushbu tadqiqot kompyuter tarmoqlarida zararli trafikni aniqlash bo‘yicha ilmiy-amaliy ahamiyatga ega bo‘lgan, zamonaviy xavfsizlik muammolariga mos, kompleks, mustahkam va yuqori aniqlikka ega intellektual modelni taqdim etadi.

ADABIYOTLAR

1. Moustafa, N., & Slay, J. (2016). UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network Data Set). *Military Communications and Information Systems Conference (MilCIS), IEEE.*
2. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward Generating Realistic Intrusion Dataset and Intrusion Traffic Characterization. *ICISSP.*
3. Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks, 51(12), 3448–3470.*
4. Lakhina, A., Crovella, M., & Diot, C. (2005). Mining Anomalies Using Traffic Feature Distributions. *ACM SIGCOMM.*
5. Sommer, R., & Paxson, V. (2010). Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. *IEEE Symposium on Security and Privacy.*
6. Buczak, A. L., & Guven, E. (2016). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys & Tutorials.*
7. Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *Computers & Security, 28(1–2), 18–28.*
8. Mishra, P., Varadharajan, V., Tupakula, U., & Pilli, E. S. (2018). A detailed investigation and analysis of using machine learning techniques for intrusion detection. *IEEE Communications Surveys & Tutorials.*
9. Aldwairi, M., & Al-Fayez, F. (2017). Deep learning and feature engineering for intrusion detection systems. *IEEE Security Conference.*
10. Zhong, Y., Guo, Q., & Wang, J. (2020). An Efficient Network Intrusion Detection Based on SVM. *Journal of Ambient Intelligence and Humanized Computing.*
11. Breiman, L. (2001). Random Forests. *Machine Learning, 45(1), 5–32.*

ISSN 3030-3907

12. Kim, H., & Kim, J. (2014). Flow-Based Malicious Traffic Detection Using Machine Learning. *International Journal of Information Security Science.*
13. Iglesias, F., & Zseby, T. (2015). Analysis of network traffic features for anomaly detection. *Machine Learning, 101(1), 59–84.*
14. Alshamrani, A., Myneni, S., Chowdhary, A., & Huang, D. (2019). A Survey on Advanced Persistent Threats. *IEEE Communications Surveys & Tutorials.*
15. Kumar, S., & Babu, M. S. (2021). Machine Learning-Based Hybrid Model for Intrusion Detection. *Journal of Information Security and Applications.*
16. Zhang, Y., Chen, X., & Xiang, Y. (2013). SVM-based effective network traffic classification. *Computer Communications, Elsevier.*
17. Kang, M. J., & Kang, J. W. (2016). Intrusion Detection System Using Deep Neural Networks. *PLOS ONE.*
18. Li, W., Meng, W., & Kwok, L. (2021). Survey on Machine Learning for Data Streams. *ACM Computing Surveys.*
19. Cheng, J., Li, Q., & Zhao, X. (2017). Real-time DDoS Detection Using Flow Features and SVM. *Security and Communication Networks.*
20. Scikit-learn Developers (2024). Scikit-learn Documentation: Machine Learning in Python. *https://scikit-learn.org*
21. Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A Detailed Analysis of the KDD CUP 99 Data Set. *IEEE Symposium on Computational Intelligence.*
22. Moore, A., & Zuev, D. (2005). Internet Traffic Classification Using ML. *ACM SIGMETRICS.*
23. Shiravi, A., Shiravi, H., Tavallaee, M., & Ghorbani, A. A. (2012). A Systematic Evaluation Framework for IDS Dataset. *IEEE Transactions on SDN.*
24. Shone, N., Ngoc, T., Phai, V., & Shi, Q. (2018). A Deep Learning Approach for Network Intrusion Detection. *IEEE Access.*
25. Yu, S., Zhou, W., Doss, R., & Jia, W. (2013). A Survey on DDoS Attacks. *IEEE Communications Surveys.*
26. Carter, K., & Dey, R. (2021). Machine Learning Approaches for Traffic Filtering. *Springer Cybersecurity.*
27. Diro, A. A., & Chilamkurti, N. (2018). DL-based IDS for IoT. *IEEE Transactions on Industrial Informatics.*
28. Verma, A., & Ranga, V. (2019). Evaluation of ML Classifiers for IDS. *Procedia Computer Science.*
29. Abaid, A., & Alazab, M. (2021). Botnet Detection in Encrypted Traffic. *Future Generation Computer Systems.*
30. McHugh, J. (2001). Testing Intrusion Detection Systems: DARPA Critique. *ACM Transactions on Information and System Security.*

ISSN 3030-3907

31. Fröhlich, P., & Dorfer, T. (2020). NetFlow Traffic Analysis Using ML. *Elsevier Computer Communications.*
32. Maiorca, D., Biggio, B., & Giacinto, G. (2019). Adversarial ML in IDS. *IEEE Transactions on Dependable and Secure Computing.*
33. Hoque, N., et al. (2012). Review of Feature Selection Methods for IDS. *International Journal of Computer Applications.*
34. Zuech, R., Khoshgoftaar, T., & Wald, R. (2015). Cyber Attack Categorization Using Data Mining. *Elsevier Journal of Big Data.*
35. Kwon, D., et al. (2019). Encrypted Traffic Growth and IDS Challenges. *IEEE Communications Surveys.*
36. Tama, B. A., & Comuzzi, M. (2019). Flow-Based Traffic Classification for Big Data. *Neurocomputing.*
37. Yin, C., Zhu, Y., Fei, J., & He, X. (2017). LSTM-based IDS. *IEEE Access.*
38. Chowdhury, M., et al. (2021). Hybrid CNN-LSTM for Anomaly Detection. *Journal of Information Security.*
39. Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2014). Network Traffic Classification and IDS Survey. *IEEE Communications Surveys.*
40. Boston, N., & Javed, A. (2020). Feature Engineering Techniques for Traffic Analysis. *Elsevier Expert Systems with Applications.*

ISSN 3030-3907