TRANSPORT POG‘ONA PROTOKOLLARI ASOSIDA SHIFRLANGAN TRAFIKNI TASNIFLASH VA BOSHQARISH MODELI

Tojiyeva Feruza Qobiljon qizi

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti doktaranti
E-mail: 11feruza@gmail.com
ORCID: 0009-0005-1261-0505

K E Y W O R D S

tarmoq trafigi, paket, protokol, shifrlangan trafik, transport pog‘onasi, xavfsizlik, tarmoq monitoring.

A B S T R A C T

Ushbu maqolada transport pog‘onasi protokollari asosida shifrlangan tarmoq trafikini tasniflash va boshqarish modeli taklif etilgan. Hozirgi kunda shifrlangan trafikning keng tarqalganligi tufayli, trafikni tasniflash va boshqarilishi masalasi dolzarb muammo bo‘lib qolmoqda. Maqolada taklif etilgan yondashuv transport pog‘onasi darajasida protokollarni tahlil qilishga asoslanadi va shifrlangan trafikni tasniflash uchun transport pog‘onasi segmentlarini ishlatadi, bu esa ilgari mavjud bo‘lgan metodlarga qaraganda samarali va tezkor yechimlar yaratadi. Ushbu dasturiy ta'minot yordamida ilovalar darajasidagi protokollarni aniqlashda 98-99% aniqlikka erishilgan. Ushbu model shifrlangan trafikni aniqlashda qo‘llanilishi mumkin bo‘lgan yangi yondashuvlarni taqdim etadi. Bunga qo‘shimcha ravishda, tarmoq trafigini boshqarish mexanizmlari ishlab chiqilib, foydalanuvchi interfeysi orqali oqimlarni nazorat qilish va cheklash imkoniyati yaratildi. Ushbu model tarmoq xavfsizligini ta'minlashda samarali vosita bo‘lib, uning amaliy qo‘llanishi tarmoq infratuzilmasining xavfsizligini yaxshilashga yordam beradi.

Kirish. Bugungi kunda internet xizmatlariga bo‘lgan ehtiyoj kundan - kun ortib bormoqda. Ijtimoiy tarmoqlarda ma’lumotlar almashinuvidan tortib, global tarmoqlarda kerakli ma’lumotlarni qidirishdagi barcha amallar internet tarmog‘i bilan bog‘liqdir. Youtube, Netflix va MegaUpload kabi ilovalar internet tarmoq tuzilishini butunlay o‘zgartirdi. Shuningdek, ITU (International Telecommunication Union) ning 2024 yilgi statistik tadqiqotlariga ko‘ra, internet foydalanuvchilarining soni 5.5 milliardga yetdi, ya'ni dunyo aholisining 68 foizini tashkil qiladi. Bu ko‘rsatgich 2023 yilga nisbatan 3 foizga, 2022 yil 2023 yilga nisbatan esa 3.5 foizga yuqorilagan [1].

Trafikni tasniflash tarmoq unumdorligini oshirishda muhim va dolzarb omil hisoblanadi. Zamonaviy tarmoq texnologiyalarini DPI (Deep Packet Inspection) texnologiyasiz tasavvur qilish qiyin, chunki DPI yordamida tarmoq hujumlarini aniqlash, korporativ tarmoqlar xavfsizligini ta'minlash va aloqa operatorlari tomonidan foydalanuvchi trafikini bloklash kabi amallarni samarali bajarish mumkin.

DPI o‘zining samaradorligiga qaramasdan, bir qancha kamchiliklarga ham ega. DPI texnologiyasi, tahlil qilinayotgan paketning foydali yuklamasini tekshirish orqali qaror qabul qiladi. Biroq, bugungi kunda shifrlangan trafiklar tarmoqning katta ulushini tashkil etadi. Shu sababli tarmoqdan o‘tayotgan paketlarni tahlil qilishda paket foydali yuklamasini tasniflash murakkablashmoqda [2,10].

Nazariy qism. Ushbu muammolarni hal etish maqsadida, DPI texnologiyasining asosiy vazifalaridan biri bo‘lgan — amaliy pog‘ona protokolini aniqlashga alternativ uslub ishlab chiqildi. Ushbu uslubni amalga oshirish uchun juda kam hajmdagi ma'lumot yetarli bo‘lib, barcha ma'lum bo‘lgan portlar ro‘yxatidan foydalanish

(cc) BY © *Tojiyeva Feruza Qobiljon qizi*

yoki paketni foydali yuklamasini tekshirishga ehtiyoj qolmaydi.

DPI texnologiyasida tasniflash obekti trafik oqimini transport pog‘onasining IP paketlari to‘plami hisoblanadi. Bunda bir xil transport pog‘onasi protokoliga ega bo‘lgan IP paketlarda o‘zaro juftliklar yaratiladi:

< (manba IP, manba port), (qabul qiluvchi IP, qabul qiluvchi port)>.

Ushbu tadqiqotda turli protokollardan foydalanadigan ilovalar, transport pog‘onasida ham turli statik xususiyatlarga ega bo‘lgan oqimlarni qanday yuzaga keltirishi o‘rganildi. Oqimlarning statik xarakteristikalarini hisoblash orqali, bu o‘lchovlar yordamida oqimni qaysi ilova hosil qilganligini va unga mos ravishda ilova pog‘onasida ishlatilgan protokol turini aniqlash mumkinligi ko‘rsatilgan.

TCP aloqasi o‘rnatilgandan keyin mijoz foydali yuklamani klient tomondagi ma'lumot bo‘laklari bo‘yicha ketma-ket uzatishni boshlaydi. Server esa foydali yuklama yubormasdan faqat tasdiqlash ACK yuboradi va mijoz tomoni ma'lumotlarni uzatishda davom etadi. Agar serverda foydali yuklamani uzatish zarurati tug‘ilsa, mijoz tomoni ma'lumot uzatishni to‘xtatadi va server ma'lumotlar bo‘lagini uzatishni boshlaydi. Bu jarayonda foydali yuklama bo‘lgan ma'lumotlar bir tomondan ikkinchi tomonga ketma-ket uzatiladi. Agar har ikkala tomonda ma'lumotlar uzatish tezlashib ketgan bo‘lsa, ma'lumot qismlari alohida paketlarga ajratilishi mumkin [7].

Ushbu ma'lumotlarga asoslangan holda biz quyidagi statik o‘lchovlarni oldik:

- Mijoz tomonidan yuborilgan, transport pog‘onasidagi segmentlar ketma-ketligi uzunligi;
- Server tomonidan yuborilgan, transport pog‘onasidagi segmentlar ketma-ketligi uzunligi;
- Mijoz tomonidan yuborilgan, transport pog‘onasidagi ma'lumotlar ketma-ketligi uzunligi;
- Server tomonidan yuborilgan, transport pog‘onasidagi ma'lumotlar ketma-ketligi uzunligi;

Yuqoridagi o‘lchovlar asosida 32 ta oqim statik xarakteristikasi hosil qilindi. Tadqiqotlar[8] shuni ko‘rsatdiki, oqimning boshlang‘ich bir nechta paketlari foydalanuvchi ilova protokoli turi bo‘yicha malumotlarni tashishi mumkin. Shu sababli, transport pog‘onasida har bir oqimning N segmentigacha kuzatish orqali yuqorida keltirilgan statik o‘lchovlarni faqatgina dastlabki N ta segmentdan olindi.

Har bir obyektda olingan xarakteristikalar mashina o‘rganishning Tasodifiy o‘rmon (Random Forest) algoritmi yordamida tasniflandi. Ushbu algoritm ko‘plab tadqiqotlarda aynan trafikni tasniflashda yaxshi natija olingan [3,4,5,6].

Boshqa tarmoq tasniflash texnologiyalari bilan solishtirilganda, tasodifiy o‘rmon algoritmi quyidagi afzalliklar bilan ajralib turadi:

- Tasniflash masalasida ishlatiladigan ilovalarda, tasodifiy o‘rmon algoritmi to‘lib ketishni oldini oladi;
- Tasniflash va regressiya masalalarida bittalik tasodifiy o‘rmon algoritmi ishlatiladi;
- Tasodifiy o‘rmon algoritmi o‘rgatish ma'lumotlar to‘plamidan eng muhim xususiyatlarni identifikatsiya qilishda ishlatilishi mumkin;

Tadqiqot usullari. Tarmoq trafikini tasniflash va boshqarish dasturiy ta'minotini ishlab chiqishda real tarmoqdan paketlarni to‘plash va ularni tahlil qilish uchun PCAP (Packet Capture) fayllari ishlatildi. Ushbu paketlar yordamida tarmoq trafigining turli xususiyatlari to‘plandi va tahlil qilindi. Tarmoq paketlarini yig‘ish uchun Python dasturlash tili va uning Scapy kutubxonasi asosida ishlab chiqilgan “Packet Saver” dasturi ishlatildi. Ushbu dastur tarmoqdan to‘plangan paketlarni real vaqt rejimida saqlash va keyinchalik tahlil qilish imkonini beradi.

Tasniflash va boshqarish jarayonida olingan traffic.pcap fayllari nDPI (Network Deep Packet Inspection) yordamida oqimlarga ajratildi. Bu jarayonning asosiy maqsadi — paketlarni to‘plab, ularni tarmoq oqimlari bo‘yicha tasniflashdir. Tasniflashni amalga oshirish uchun Python dasturlash tilining Sklearn kutubxonasi yordamida Random Forest algoritmining o‘rganish modeli yaratildi. Tarmoq trafigini tasniflash sohasida olib

(cc) BY © Tojiyeva Feruza Qobiljon qizi

borilgan tadqiqotlar o‘rganilib, tasodifiy o‘rmon algoritmi tarmoq trafigini tasniflashda samarali ekanligi isbotlangan[3]. Shu sababli, ushbu tadqiqot aynan shu algoritm asosida olib borildi.

Model Maker dasturi yordamida bu jarayon avtomatlashtirilgan bo‘lib, u quyidagi ketma-ketlikda ishlaydi:

1. PCAP fayli yig‘iladi va nDPI yordamida oqimlarga ajratiladi.
2. Olingan oqimlar Sklearn kutubxonasi yordamida tasniflanadi.
3. Tasodifiy o‘rmon algoritmi yordamida o‘rganish modeli yaratilib, tarmoq trafigi protokollarini aniqlash jarayoni amalga oshiriladi.

Tasniflashning ikki bosqichli jarayoni: Birinchi bosqichda taqdim etilgan ma'lumotlar asosida daraxtlar qurilishi amalga oshiriladi. Bu jarayon statistik tahlilni talab qiladi, chunki daraxtlar hosil qilish uchun har bir oqimning segmentlari va ularning xususiyatlari hisobga olinadi. Ushbu jarayonda, oqimni tasniflashda qaysi statistika ma'lumotlarining muhim rol o‘ynashini aniqlash uchun foizlar yordamida tahlil qilish amalga oshiriladi.

Ikkinchi bosqichda esa hosil qilingan daraxtlar asosida tekshiruv o‘tkaziladi va tasniflash natijalari tekshiriladi. Bu jarayon davomida protokollarni aniqlash uchun statistik tahlil qilishning ahamiyati katta.

Tadqiqot natijalari. Olingan natijalar, ya'ni protokollarni to‘g‘ri aniqlash darajasi 1-rasmda ifodalangan.

Tarmoq tahlilining samaradorligini oshirish uchun foizli qiymatlar, protokollarni aniqlashda foydalanilgan xususiyatlar va ular orasidagi bog‘liqlik haqida batafsil ma'lumot beriladi. Natijada, tasodifiy o‘rmon algoritmi yordamida protokollarni aniqlashda aniq va samarali natijalarga erishish mumkin.

	precision	recall	f1-score	support
BitTorrent	0.77	0.97	0.86	62
DNS	1.00	0.99	1.00	4338
HTTP	0.97	1.00	0.98	1750
Quic	1.00	1.00	1.00	74
SSL	0.96	0.90	0.93	760
avg / total	0.99	0.98	0.98	6984

	BitTorrent	DNS	HTTP	Quic	SSL
BitTorrent	60	0	0	0	2
DNS	0	4314	0	0	24
HTTP	1	0	1746	0	3
Quic	0	0	0	74	0
SSL	17	2	59	0	682

Rasm 1. *Tarmoq trafikining protokol tasniflash natijalari: Precision, Recall, F1-score va Confusion Matrix*

Natijalar shuni ko‘rsatdiki, aniqlashdagi eng kam ko‘rsatkich BitTorrent trafigi bo‘lib, 77% ni tashkil qilgan va testlash uchun 62 ta BitTorrent oqimini ishlatgan. Eng yuqori aniqlik esa DNS va QUIC protokollarida kuzatilgan. Bu protokollar bo‘yicha aniqlik 100% ga erishilgan. Pastdagi

(cc) BY © Tojiyeva Feruza Qobiljon qizi

qismda esa protokollarni bir-biriga almashtirish ehtimoli ko‘rsatilgan.

1-jadval.

O‘rganish va testlash uchun ketgan paketlar soni

Protokol	O‘rganishda ishlatilgan oqimlar soni	Testlashda ishlatilgan oqimlar soni
DNS	2169	4338
HTTP	875	1751
SSL	291	760
Quic	37	74
BitTorrent	31	62

Yuqaridagi jadvaldan ko‘rinib turibdiki, tasodifiy o‘rmon algoritmi uchun testlash uchun sarflangan oqimlar soni o‘rganish uchun sarflangan oqimlar sonidan deyarli ikki barobar ko‘p. Bu esa testlash uchun ko‘proq ma'lumot olinishi va natijada o‘rganish darajasini aniq hisoblab chiqish imkonini beradi.

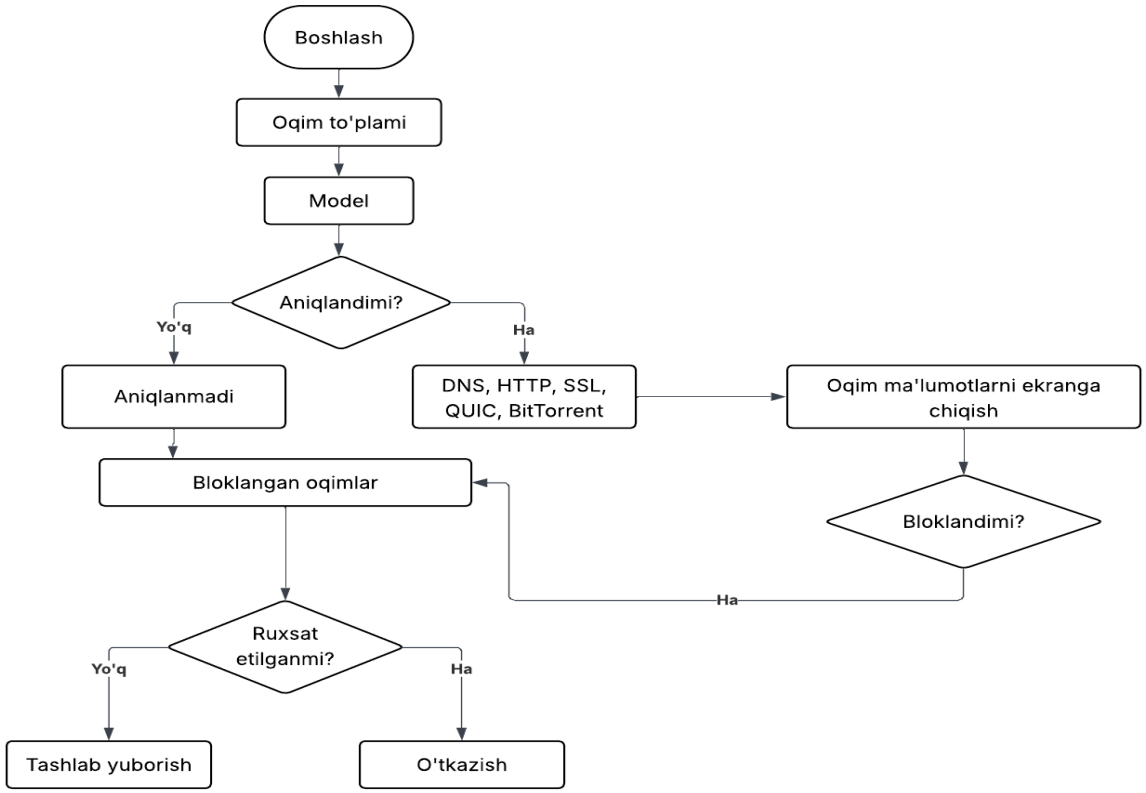
Quyidagi 2-jadvaldan ko‘rinib turibdiki, protokollarni aniqlash imkoniyati eng kichik qiymati 98% ni tashkil etadi, bu BitTorrent trafigi uchun tegishli. Qolgan turlar esa deyarli 100% aniqlik bilan topilgan.Ushbu ko‘rsatkichlardan kelib chiqib, biz tanlagan usul va algoritmning aniqligi 99% ni tashkil etadi, bu esa dasturimizning aniq ishlayotganini isbotlaydi.

2-jadval.

Ilova protokollarining aniqlanish ko‘rsatgichi

Protokol turi	Aniqliligi
DNS	1.00
HTTP	0.99
SSL	1.00
Quic	1.00
BitTorrent	0.98

Quyidagi 2-rasmda tarmoq trafikini tasniflash va boshqarish dasturiy ta'minotining ishlash algoritmi.



Rasm 2. *Tarmoq trafikini tasniflash va boshqarish dasturiy ta'minotining ishlash algoritmi*

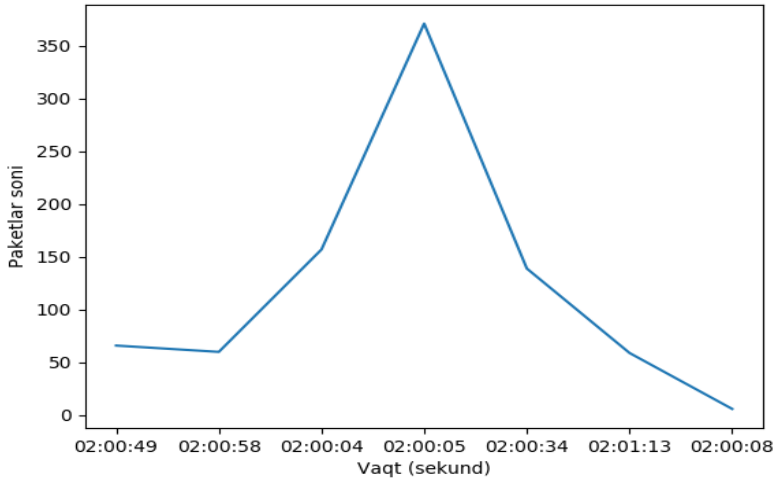
© Tojiyeva Feruza Qobiljon qizi

Model orqali tasniflangan paketlar oqimi tekshirish qismiga va foydalanuvchi interfeysi orqali yuboriladi. Agar oqimga hech qanday cheklov qo‘yilmagan bo‘lsa, u tarmoqka kiradi yoki chiqib ketadi. Agar oqimga cheklovlar qo‘yilgan bo‘lsa, bunday paketlar rad etiladi.

Foydalanuvchi interfeysi orqali, tarmoq oqimlarini nazorat qilish mumkin. Foydalanuvchi, nazorat qilinayotgan oqimlardan birini bloklash uchun uni belgilab, unga cheklov qo‘yishi mumkin. Shu tarzda, foydalanuvchi o‘zining xohishiga ko‘ra, tarmoqdagi oqimlarni boshqarishi mumkin.

Boshqarish jarayoni oqim paketlarini o‘tkazish yoki tashlab yuborish orqali amalga oshiriladi. Bu jarayon Python dasturlash tilining os moduli yordamida boshqariladi.

Bundan tashqari, foydalanuvchi interfeysi orqali o‘tayotgan tarmoq trafikining vaqt bo‘yicha paketlar soni aniqlanadi. Bu ma'lumotlar orqali tarmoqning yuklanishi va uning dinamikasi haqida umumiy tasavvur olish mumkin. 3-rasmda tarmoq trafigining vaqt bo‘yicha soni qanday o‘zgarishini ko‘rsatadigan grafik mavjud.



Rasm 3. *Umumiy paketlar sonini vaqtga bog‘liqlik grafigi*

Ushbu grafikda tarmoq trafigining vaqt o‘tish bilan qanday o‘zgarishi ko‘rsatilgan. Paketlar soni va vaqt (sekundlarda) o‘rtasidagi bog‘lanish tasvirlangan. Bu yerda: Vaqt (sekundlarda) - bu o‘rinda tarmoqda ma'lum bir vaqt oralig‘ida sodir bo‘lgan hodisalarni ko‘rsatadi. Paketlar soni - bu esa vaqt davomida o‘tgan paketlarning sonini ifodalaydi.

Vaqtning biron-bir nuqtasida paketlar sonining keskin o‘sishi yoki kamayishi tarmoqda yuqori yuklanish yoki tarmoq trafigining o‘zgarishini anglatadi.

Natijalar va muhokama. Tarmoq trafigini tasniflashda Random Forest algoritmi yordamida olingan natijalar yuqori samaradorlikni ko‘rsatdi. DNS va QUIC protokollari 100% aniqlik bilan aniqlangani, ushbu protokollar uchun algoritmning yuqori darajada samarali ishlashini ta'kidlaydi. BitTorrent protokoli esa 77% aniqlik bilan aniqlangan bo‘lib, bu yuqori darajada noaniqlikni ko‘rsatadi. Ammo, bu natija BitTorrent trafigining o‘ziga xos xususiyatlari bilan bog‘liq bo‘lishi mumkin. Ushbu natijalar Moore va Papagiannaki [8] tomonidan o‘tkazilgan tadqiqotlarga o‘xshash bo‘lib, ular ham protokollarni aniqlashda aniqlikning farqli darajalarini ko‘rsatganlar.

Algoritmning ishlash samaradorligi 98-99% aniqlik bilan ko‘rsatdi, bu esa Random Forest algoritmining tarmoq trafigini tasniflashda samarali ekanligini tasdiqlaydi. Bundan tashqari, testlash uchun sarflangan oqimlar soni o‘rganish uchun sarflangan oqimlar sonidan deyarli ikki barobar ko‘p bo‘lishi, ma'lumotlar hajmining ortishi algoritmning aniqligini oshirishga yordam berganligini ko‘rsatadi.

Kelgusida, boshqa metodlar va algoritmlar bilan solishtirganda Random Forest algoritmining afzalliklarini o‘rganish zarur. Xususan,

(cc) BY © Tojiyeva Feruza Qobiljon qizi

shifrlangan trafikni aniqlashda yangi yondashuvlarni ishlab chiqish va tarmoqning yuqori yuklanishi holatida tizimning samaradorligini oshirish uchun takliflar kiritish kerak.

Xulosa. Ushbu maqolada tarmoq trafikini transport pog‘onasi protokoli asosida aniqlash va tasniflash innovatsion usuli ko‘rib chiqildi. Yondashuvning asosiy maqsadi – transport pog‘onasi darajasida tahlil qilinadigan bir nechta segmentlar yordamida ilovalar darajasidagi protokolni aniqlashdir. Ushbu metod yordamida shifrlangan trafikni tahlil qilish imkoniyati yuzaga keldi. Transport pog‘onasi segmentlaridan foydalanib, shifrlangan paketlarni aniqlash va tasniflash mumkin bo‘ldi, bu esa tarmoq xavfsizligini oshirishda yangi imkoniyatlar yaratadi.

Tadqiqot jarayonida, birinchi bosqichda transport pog‘onasi protokoli asosida statistik ma'lumotlar yig‘ilib, ilova protokoli aniqlandi. Ushbu jarayon, shuningdek, shifrlangan trafikni tasniflash uchun zarur bo‘lgan imkoniyatlarni taqdim etdi. Natijada, shifrlangan paketlarni tahlil qilmasdan, faqat transport pog‘onasi darajasida ilova protokolini aniqlash mumkin bo‘ldi.

Keyingi bosqichda ishlab chiqilgan dasturiy ta'minot samaradorligini baholash uchun turli testlar o‘tkazildi. Dastur yordamida ilova protokollarini aniqlashda 98-99% aniqlik ko‘rsatildi, bu esa dasturimizning yuqori samaradorlik bilan ishlashini ko‘rsatadi. Testlash natijalari, dastur yordamida protokollarni to‘g‘ri aniqlash darajasini va uning boshqa ilovalar bilan o‘xshashligini aniq ko‘rsatdi. Ikkinchi bosqichda esa, tarmoq oqimlarini boshqarish masalasi ko‘rib chiqildi.

Dasturda tarmoqka ruxsat etilmagan paketlar aniqlanib, ularning tarmoqka kirishiga yo‘l qo‘yilmagan. Oqimlarni boshqarish va tasniflash jarayonlari avtomatik tarzda amalga oshirildi, bu esa tizimni samarali boshqarish va xavfsizlikni ta'minlashga yordam berdi.

Natijada, ishlab chiqilgan metod va dasturiy ta'minot tarmoq trafigini tasniflashda yuqori samaradorlikni ta'minladi va kelajakda tarmoq xavfsizligini oshirish va boshqarishni yanada takomillashtirish uchun muhim qadam bo‘lib xizmat qiladi.

Adabiyotlar ro‘yxati

1. Petrosyan, A. Worldwide digital population 2025 // Statista. — Apr 1, 2025
2. Shafiq, M., Yu, X., Laghari, A.A., Yao, L., Karn, N.K., Abdessamia, F. Network Traffic Classification Techniques and Comparative Analysis Using Machine Learning Algorithms // In: *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, China, pp. 2451-2455. — 2016.
3. Azab, A., Khasawneh, M., Alrabaee, S., Choo, K.-K. R., & Sarsour, M. Network traffic classification: Techniques, datasets, and challenges // *Digital Communications and Networks*, 8(1), 1-11. — 2022.
4. Li, W., & Moore, A. W. A Machine Learning Approach for Efficient Traffic Classification // IEEE Xplore. — 2021.
5. Mondal, P. K., Aguirre, S. L., Benedetto, E., Shen, Y., Guo, M. A dynamic network traffic classifier using supervised ML for a Docker-based SDN network // *Connection Science*, 28(4), 293-308. — 2021.
6. Zhao, J., Jing, X., Yan, Z., Pedrycz, W. Network traffic classification for data fusion: A survey // *Information Fusion*, 72, 22-47. — 2021.
7. Bernaille, L., Teixeira, R., Salamatian, K. Early Application Identification // *ACM CoNEXT 2006 Conference Proceedings*. — 2006.
8. Moore, A. W., Papagiannaki, K. Toward the Accurate Identification of Network Applications // *Proceedings of the 6th International Conference on Passive and Active Network Measurement (PAM'05)*, 41–54. — 2005. — Boston, Massachusetts, USA.
9. Rachmawati, S. M., Kim, D.-S., Lee, J. M. Machine Learning Algorithm in Network Traffic Classification // In: *2021 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 320–325. — IEEE, Jeju Island, 2021.
10. Lim, H.-kyo, Kim, J.-Bong, Han, Y.-H. Packet-based Network Traffic Classification Using Deep Learning // In: *Digital Signal Processing and Communications*, February, 2019.

(cc) BY © *Tojiyeva Feruza Qobiljon qizi*