DASTURIY ANIQLANGAN TARMOQLARDA LOGLARNI SUN’IY INTELLEKTLI TAHLILLASH ORQALI ANOMALIYALARNI ANIQLASH USULI

Jumayev Sodikjon Nuraliyevich[1], Bozorov Suhrob[2]
[1]Denov tadbirkorlik va pedagogika instituti,
E-mail: jumayev.sodiq9091@mail.ru, ORCID: 0009-0001-4741-8848
[2]Muhammad al-Xorazmiy nomidagi TATU,
E-mail: mr.bozorov95@gmail.com, ORCID: 0000-0002-2377-934X

K E Y W O R D S

Dasturiy aniqlangan tarmoqlar (SDN), anomaliyalarni aniqlash, log fayllarini tahlil qilish, sun’iy intellekt, mashina o‘qishi, kiberxavfsizlik, tarmoq xavfsizligi, K-means, PCA, Decision Tree.

A B S T R A C T

Zamonaviy axborot-kommunikatsiya tizimlarining rivojlanishi bilan dasturiy aniqlangan tarmoqlar (SDN) keng qo‘llanilmoqda. SDN tarmoqlarining markazlashtirilgan boshqaruv imkoniyati tarmoq infratuzilmasini moslashuvchan va dinamik qiladi, biroq bu xavfsizlik jihatidan yangi tahdidlarni ham keltirib chiqaradi. Ushbu maqolada SDN tarmoqlarda vujudga keladigan xavfsizlik tahdidlari, jumladan, konfiguratsiya xatoliklari, boshqaruv mexanizmining ochiqligi, avtorizatsiya zaifliklari va tarmoqni manipulyatsiya qilish hujumlari tahlil qilingan. Bu tahdidlarni aniqlash va ularga qarshi chora ko‘rish uchun log fayllarini tahlil qilish samarali vosita sifatida qaraladi. Loglarni tahlil qilish orqali anomaliyalarni aniqlashda sun’iy intellekt, mashina o‘qishi algoritmlari (masalan, K-means, PCA, Decision Tree, Isolation Forest) keng qo‘llanilmoqda. Maqolada loglarni tahlil qilishning tizimli usullari — loglarni yig‘ish, strukturalash, xususiyatlarni ajratib olish va anomaliyalarni aniqlash jarayonlari keltirilgan. Eksperimental natijalar log tahlilida sun’iy intellektli yondashuvlarning anomaliyalarni aniqlashda yuqori aniqlik (0.986) va F1-qiymat (0.710) ko‘rsatayotganini tasdiqlaydi.

Kirish

Zamonaviy axborot-kommunikatsiya tizimlarining asosiy qismlaridan biri bu– kompyuter tarmoqlaridir. Ular o‘zaro bog‘langan kompyuterlar va boshqa qurilmalar orqali ma’lumotlarni tezkor va ishonchli almashish imkonini yaratadi. Kompyuter tarmoqlari nafaqat texnologik infratuzilma sifatida, balki ma’lumotlarni uzatish, qayta ishlash va boshqarishning asosiy vositasi sifatida ham muhim o‘rin tutadi.

SDN (dasturiy aniqlangan tarmoqlar) arxitekturasi esa, ma’lumotlarni uzatuvchi va boshqaruvchi qatlamlarni ajratishga asoslanadi. Bu yondashuvda, tarmoqning boshqaruvi markaziy tizim tomonidan amalga oshiriladi, lekin ma’lumotlar uzatiladigan qatlamda ma’lumotlarni uzatish jarayoni mustaqil bo‘ladi. Bunday ajratish dinamik boshqaruv va real vaqtli monitoring imkoniyatlarini taqdim etadi, bu esa tizimning moslashuvchanligini oshiradi. Ammo, tizimning murakkabligi va resurslarga yuqori talablar xavfsizlikni boshqarishda yangi qiyinchiliklar yaratadi. Shuningdek, yuklab olish va ishga tushirishda tarmoq kechikishlari paydo bo‘lishi mumkin.

SDN tarmoqlarining eng asosiy xususiyati tarmoq boshqaruvining markazlashtirilganligi va dasturiy nazorat ostida bo‘lishidir. Bu, tarmoq administratorlariga butun tarmoqni bir nechta markaziy boshqaruv nuqtalaridan nazorat qilish imkoniyatini beradi. Shuningdek, SDN tizimlarining bir nechta muhim afzalliklari mavjud:

- Moslashuvchanlik: SDN texnologiyasi tarmoqni tez va samarali sozlash imkoniyatini

beradi. Tarmoq resurslarini markazlashtirilgan tarzda boshqarish, kerakli o‘zgarishlarni kiritishni osonlashtiradi va tarmoqni tezda qayta konfiguratsiya qilish imkoniyatini yaratadi.

- Soddalashgan boshqaruv: Tarmoqning barcha elementlarini bir markazdan boshqarish tarmoqni konfiguratsiya qilish, monitoring qilish va xavfsizlikni ta’minlashni osonlashtiradi. SDN tarmoqlarida dasturiy boshqaruvning mavjudligi tizimni boshqarishni soddalashtiradi va tezroq amalga oshirilishini ta’minlaydi.
- Kengaytirilgan xavfsizlik: SDN tizimlari tarmoq xavfsizligini markazlashtirilgan tarzda boshqarishga imkon beradi. Tarmoqning boshqaruv mexanizmlari orqali xavfsizlik devorlarini va filtrlashni samarali boshqarish mumkin. Bu, ayniqsa, tarmoqdagi noxush holatlarni tezda aniqlash va ularga qarshi chora ko‘rishni osonlashtiradi.
- Tezkor innovatsiyalar: SDN tizimlari yangi texnologiyalarni qo‘llashni osonlashtiradi. Tarmoqning dasturiy boshqaruvi yordamida yangi protokollarni qo‘llash va tarmoqni yangilash jarayonlari tez va samarali amalga oshiriladi.

SDN arxitekturasida trafikni analiz qilish va tarmoqni himoya qilish uchun o‘ziga xos algoritmlar va vositalar ishlab chiqilgan. Misol uchun, Ghosh va boshqalar SDN yordamida tarmoqda anomaliyalari aniqlashning samaradorligini o‘rganishgan. Tadqiqotda SDN arxitekturasida tarmoqni himoya qilishning yangi usullari tavsiya etilgan va tarmoqda yuzaga keladigan zararli faoliyatlarni aniqlashda SDN tizimining ahamiyati tasdiqlangan [8].

SDN tarmoqlarda amalga oshiriluvchi tahdidlar va ularni oldini olish

SDN tarmoqlarda bir qancha xavfsizlik muammolari ham mavjud bo‘lib, ushbu tahdidlar tashkilot va foydalanuvchilar ma’lumotlariga jiddiy xavf tug‘diradi.

1-jadval.

SDN tarmoqlarda tahdidlar tasnifi

Tahdid turi	Tavsif	SDN tarmoqlariga ta’siri
Sozlamalar va konfiguratsiya xatoliklari	SDN tarmog‘idagi konfiguratsiya xatoliklari yoki noto‘g‘ri sozlamalar hujumchilarga tarmoqni manipulyatsiya qilish imkonini beradi.	Noto‘g‘ri konfiguratsiya tarmoqning ishlashini buzishi va xavfsizlikni zaiflashtirishi mumkin.
Tarmoqning ochiqligi va boshqaruv mexanizmi	SDN tarmoqlarida boshqaruv mexanizmining ochiqligi, hujumchilarga tarmoqdagi barcha jarayonlarni o‘zgartirish imkoniyatini beradi.	Boshqaruv mexanizmiga kirish orqali tarmoqni to‘liq boshqarish, ma’lumotlarni yo‘naltirishni buzish yoki o‘g‘irlikka olib kelishi mumkin.
Avtorizatsiya va kirish nazorati zaifliklari	Tarmoqning markazlashtirilgan boshqaruvi orqali foydalanuvchi kirish huquqlarini noto‘g‘ri tartibga solish, hujumchilarga tizimga kirishga imkon beradi.	Ruxsat etilmagan foydalanuvchilarga tizimga kirish imkoniyatini beradi, bu esa xavfsizlikni buzadi va ma’lumotlarni o‘g‘irlashga olib keladi.
Tarmoqni manipulyatsiya qilish hujumlari	SDN tarmoqlarining markazlashtirilgan boshqaruv tizimida manipulyatsiya qilish orqali tarmoqda noto‘g‘ri marshrutlash yoki tarmoqni noqulay holatga keltirish.	Hujumchilar tarmoqni manipulyatsiya qilib, barcha trafikni o‘zgartirishi yoki tizimni offline holatga keltirishi mumkin.

Xavfsizlik protokollaridagi zaifliklar	SDN tarmoqlarida xavfsizlik protokollaridagi zaifliklar, masalan, shifrlashning zaifligi yoki ma’lumotlarni autentifikatsiya qilishdagi xatoliklar.	Tarmoqdagi ma’lumotlar va trafikning shifrlanishi yoki autentifikatsiyasi buzilishi tarmoqning xavfsizligini zaiflashtiradi.

Yuqoridagi jadvalda keltirilgan tahdidlar uyushtirilganda SDN tarmoqlari barja harakatlarni avomatik qayd etadi va bu yozuvlar log fayllariga saqlanadi. Log fayllar va ulardagi yozuvlarni tahlil qilish tizimdagi anomal holatlarni aniqlash imkoni beradi.

Log fayli yoki hodisa qayd yozuvlari - bu kompyuter “logged” voqealar ro‘yxatini o‘z ichiga olgan fayl. Ular tizim bilan bog‘liq ma’lumotlarni, shu jumladan internetdan foydalanishni o‘z ichiga olgan yozuvlardir. Hodisa qayd yozuvlari ishlaydigan ilovalar, xizmatlar, tizim xatolari va yadro xabarlari haqidagi ma’lumotlarni o‘z ichiga oladi. Log fayllari odatda dasturiy ta’minotni o‘rnatish paytida va veb-serverlar tomonidan yaratiladi, lekin ular boshqa turli maqsadlarda, jumladan, kiberxavfsizlikda hujumlar va ularning anomaliyalarini aniqlash hamda tahlil qilishda ham ishlatilishi mumkin.

Log fayli to‘rtta asosiy darajasi mavjud:

1. debug - keng ko‘lamli holat o‘tishlarini qayd etish: ma’lumotlar bazalariga kirish, xizmatni ishga tushirish va to‘xtatish;
2. ogohlantirish - favqulodda vaziyatlar, masalan, noto‘g‘ri so‘rov formati;
3. xato - tipik xatolarni qayd etish;
4. fatal - halokatli ishdan chiqish: ma’lumotlar bazasiga kirish taqiqlangan, diskda bo‘sh joy etishmasligi.

Ro‘yxatga olish bo‘yicha uchta qo‘shimcha darajasi mavjud:

1. trace - jarayonni bosqichma-bosqich qayd etish;
2. fault - muammoni lokalizatsiya qilish qiyin bo‘lganda kerak;
3. service log - xizmat, xizmat faoliyati haqida umumiy ma’lumot.

Log fayli turlari:

1. Server - serverga qo‘ng‘iroqlar va qo‘ng‘iroqlar paytida yuzaga keladigan xatolar;
2. Tizim - barcha tizim hodisalari;
3. Avtorizatsiya va autentifikatsiya jurnallari - kirish va chiqish jarayonlari, kirish muammolari va boshqalar;
4. Tizimdagi ilovalar jurnallari;
5. Ma’lumotlar bazasi jurnallari - ma’lumotlar bazasiga kirish.

Log fayllarning asosiy tizimlardagi turlari

Perimetr qurilma jurnallari. Perimetr qurilmalari tarmoqqa va tarmoqdan trafikni nazorat qiladi va tartibga soladi. Xavfsizlik devorlari, virtual xususiy tarmoq (VPN), tajovuzni aniqlash tizimlari (IDS) va kirishni oldini olish tizimlari (IPS) perimetr qurilmalarining ba’zilaridir. Ushbu qurilmalar katta hajmdagi ma’lumotlarni o‘z ichiga olgan jurnallarni yaratadi va perimetri qurilma jurnallari tarmoqda sodir bo‘layotgan xavfsizlik hodisalarini tushunish uchun juda muhimdir. Syslog formatidagi jurnal ma’lumotlari AT ma’murlariga xavfsizlik tekshiruvlarini o‘tkazish, operatsion muammolarni bartaraf etish va korporativ tarmoq orqali va undan o‘tadigan trafikni yaxshiroq tushunishga yordam beradi.

Windows voqealar jurnallari. Windows OTda besh xil turdagi hodisalar ro‘yxatga olinishi mumkin. Bularning barchasida aniq belgilangan umumiy ma’lumotlar mavjud va ular ixtiyoriy ravishda hodisalarga tegishli ma’lumotlarni o‘z ichiga olishi mumkin.

Ilova biron bir hodisa haqida xabar berganida hodisa turini ko‘rsatadi. Har bir hodisa bitta turda

bo‘lishi kerak. Hodisalar jurnalini ro‘yxat ko‘rinishida har bir tur uchun boshqa belgi ko‘rsatiladi.

Quyidagi jadvalda hodisalarning ro‘yxatga olishda ishlatiladigan beshta turi tasvirlangan.

2-jadval.

Hodisalarning ro‘yxatga olishda ishlatiladigan turilari

Hodisa	Tavsifi
Xatolik	Ma’lumotni yo‘qotish yoki funksionallikni yo‘qotish kabi muhim muammoni ko‘rsatadigan voqea. Masalan, agar xizmat ishga tushirish paytida yuklana olmasa, xatolik hodisasi qayd etiladi.
Ogohlantirish	Hodisa juda ahamiyatli yemas, lekin kelajakda yuzaga kelishi mumkin bo‘lgan muammolarni ko‘rsatishi mumkin. Masalan, diskda bo‘sh joy kam bo‘lsa, ogohlantirish hodisasi qayd etiladi. Agar ilova biron bir hodisani funksional yoki ma’lumot yo‘qotmasdan tiklay olsa, u odatda voqeani ogohlantirish voqeasi deb tasniflashi mumkin.
Axborot	Ilova, drayver yoki xizmatning muvaffaqiyatli ishlashini tasvirlaydigan voqea. Masalan, tarmoq drayveri muvaffaqiyatli yuklanganda, axborot hodisalarini qayd etish o‘rinli bo‘lishi mumkin. Esda tutingki, desktop ilovalarda har safar boshlanganida biron-bir voqeani qayd etish noto‘g‘ri bo‘ladi.
Muvaffaqiyatli audit	Muvaffaqiyatli tekshirilgan xavfsizlikka kirish urinishlarini yozib oladigan hodisa. Masalan, foydalanuvchining tizimga muvaffaqiyatli urinishi muvaffaqiyatli audit hodisasi sifatida qayd etiladi.
Muvaffaqiyatsiz audit	Tekshirilgan xavfsizlikka kirish urinishining muvaffaqiyatsiz tugaganligini qayd etadigan hodisa. Masalan, agar foydalanuvchi tarmoq drayveriga kirishga harakat qilsa va muvaffaqiyatsiz bo‘lsa, urinish Muvaffaqiyatsiz audit hodisasi sifatida qayd etiladi.

Oxirgi nuqta jurnallari. Oxirgi nuqta - tarmoqdagi kompyuter, noutbuk, smartfon, server yoki ish stantsiyasi kabi hisoblash qurilmasi hisoblanadi. Oxirgi nuqtalar dasturiy ta’minot stekining turli qatlamlaridan - apparat, operatsion tizim, drayver, ma’lumotlar bazasi va ilovalardan ko‘plab jurnallarni yaratadi. Oxirgi nuqta jurnallari stekning quyi darajalaridan olinadi va oxirgi qurilmaning holati, faoliyati va samaradorlik darajasini tushunish uchun ishlatiladi.

Ilova jurnallari. Serverlar yoki oxirgi foydalanuvchi qurilmalarida ishlaydigan ilovalar voqealarni yaratadi va jurnalga kiritadi. Windows operatsion tizimi ishga tushirish, o‘chirish, yurak urishi va ishlayotgan ilovalardan ish vaqtidagi xato hodisalarini to‘playdigan markazlashtirilgan voqealar jurnalini taqdim etadi. Linuxda ilovalar jurnali xabarlarini /va /log jildida topish mumkin. Bundan tashqari, jurnal agregatorlari elektron pochta, veb-serverlar yoki ma’lumotlar bazasi serverlari kabi korporativ ilovalardan jurnallarni to‘g‘ridan-to‘g‘ri to‘plashi va tahlil qilishi mumkin.

Proksi jurnallar. Proksi-serverlar maxfiylikni ta’minlash, kirishni tartibga solish va tarmoqli kengligini tejash orqali tashkilot tarmog‘ida muhim rol o‘ynaydi. Barcha veb-so‘rovlar va javoblar proksi-server orqali o‘tganligi sababli, proksi-server jurnallari foydalanish statistikasi va oxirgi nuqta foydalanuvchilarining ko‘rish xatti-harakatlari haqida qimmatli ma’lumotlarni ochib berishi mumkin.

SDN tarmoqlarida hujumlarni aniqlashda log fayllardan foydalanish usullari

SDN tarmoqlarida hujumlarni aniqlash uchun log fayllaridan foydalanishning bir necha usullari mavjud:

1. Anomaliyalarga asoslangan aniqlash: Bu tizim xatti-harakatlaridagi anomaliyalarni aniqlash uchun joriy jurnal ma’lumotlarini tarixiy ma’lumotlar bilan solishtirishni o‘z ichiga oladi. Masalan, tizimga kirishga muvaffaqiyatsiz urinishlar sonining ko‘payishi yoki noodatiy manbadan kelgan tarmoq trafigi potentsial hujumni ko‘rsatishi mumkin.
2. Imzoga asoslangan aniqlash: Bu jurnal ma’lumotlaridagi ma’lum hujumlar yoki zararli faoliyatni aniqlash uchun oldindan belgilangan imzo yoki naqshlardan foydalanishni o‘z ichiga oladi. Misol uchun, ma’lum bir buyruqlar ketma-ketligini yoki ma’lum botnet bilan bog‘langan ma’lum IP-manzilni qidirish.
3. Log korrelyatsiyasi: Bu hujumni ko‘rsatishi mumkin bo‘lgan hodisalarning naqshlari yoki ketma-ketligini aniqlash uchun bir nechta manbalardan jurnal ma’lumotlarini tahlil qilishni o‘z ichiga oladi. Misol uchun, bir qator muvaffaqiyatsiz kirish urinishlari va undan keyin muvaffaqiyatli kirish qo‘pol kuch hujumini ko‘rsatishi mumkin.
4. Foydalanuvchilar va ob’ektlar xatti-harakatlari tahlili: Bu foydalanuvchilar va ob’ektlar uchun normal xatti-harakatlarning asosiy chizig‘ini o‘rnatish uchun jurnal ma’lumotlarini tahlil qilishni va keyin hujumni ko‘rsatishi mumkin bo‘lgan anomal xatti-harakatni aniqlash uchun mashinani o‘rganish algoritmlaridan foydalanishni o‘z ichiga oladi. Masalan, odatda ma’lum vaqtlarda tizimga kiradigan va ma’lum resurslarga kiruvchi foydalanuvchining noodatiy faoliyatini aniqlash.
5. Haqiqiy vaqtda monitoring: Bu sodir bo‘lganda shubhali faoliyat haqida ma’murlar yoki xavfsizlik xodimlarini ogohlantirish uchun real vaqtda monitoring vositalaridan foydalanishni o‘z ichiga oladi. Masalan, foydalanuvchi normal kirish darajasidan tashqaridagi manbaga kirishga harakat qilganda yoki muhim tizim sozlamalari o‘zgartirilganda ogohlantirish olish.
6. Kompyuter-ekspertiza tahlil: Bu hujum qanday amalga oshirilganligini va qanday ma’lumotlarga kirishni aniqlash uchun hujum sodir bo‘lganidan keyin jurnal ma’lumotlarini tekshirishni o‘z ichiga oladi. Masalan, qaysi tizimlar ma’lumotlar eksfiltratsiyasi hujumida ishtirok etganligini aniqlash uchun tarmoq trafik jurnallarini ko‘rib chiqish yoki buzilgan tizimga kirish uchun qaysi hisob ma’lumotlari ishlatilganligini aniqlash uchun autentifikatsiya jurnallarini tekshirish [7].

Log fayllarini shu yo‘llar bilan ishlatib, tashkilotlar xavfsizlik hodisalarini faol ravishda aniqlashlari va ularga javob berishlari, natijada ularning umumiy xavfsizlik holatini yaxshilashlari mumkin.

Log fayllarni tahlil qilish orqali anomaliyalarni aniqlash quyidagi jarayonlarni qamrab oladi:

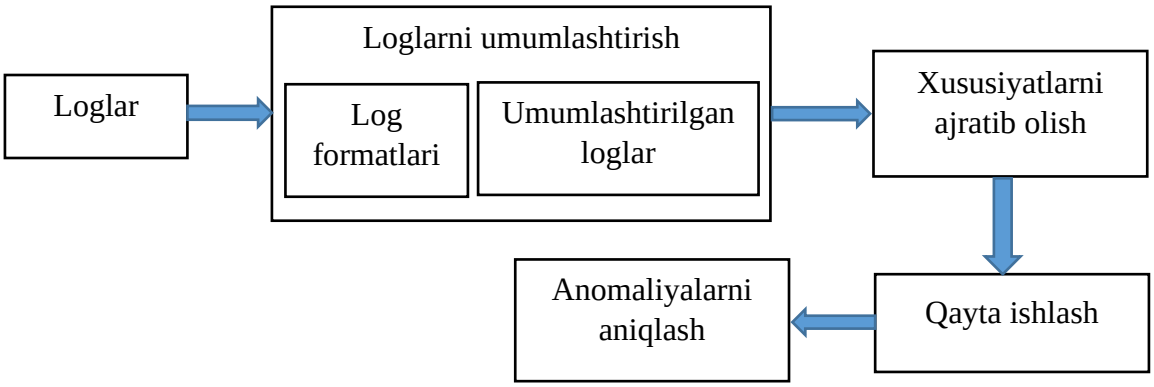
K-means algoritmi yordamida Proof-of-concept usuli

Ushbu usulning maqsadi loglarda anomaliyalarni aniqlash kontseptsiyasini (PoC) isbotlashdir. Jurnalni tahlil qilishda xavfsizlik bilan bog‘liq barcha kiruvchi harakatlarni aniqlash uchun yagona eng yaxshi yechim yo‘q. Anomaliyalarni aniqlash - bu yondashuvlardan biri bo‘lib, uning o‘zi turli xil vaziyatlar uchun turli yondashuvlar va yechimlarni o‘z ichiga oladi. Shuning uchun, PoCni amalga oshirishdan maqsad eng yaxshi natijaga ega yechim topish emas. Buning o‘rniga, maqsad bazani yaratish va anomaliyalarni aniqlash tushunchasini joriy etishdir [2].

1-rasmda PoCni amalga oshirishning umumiy ko‘rinishi keltirilgan. Testlar 64 bitli Windows operatsion tizimida, 16 Gb tezkor xotiraga ega Intel

Core i5-8400 protsessorida o‘tkazildi va kod Python 3.8 da yozilgan.

Dataset: Xavfsizlik loglarida muhim ma’lumotlar bo‘lishi mumkinligi qimmatli ma’lumotlar to‘plamini yaratish yoki topishni qiyinlashtiradi. Amalga oshirilgan modelni baholash uchun supersized ma’lumotlar bazasini topish kerak. Bunday ma’lumotlar to‘plamini yaratish ko‘p vaqt talab etadi, shuning uchun umumiy log ma’lumotlar to‘plamidan foydalaniladi. Log ma’lumotlar to‘plami He va boshqalar tomonidan taqdim etilgan HDFS ma’lumotlaridan iborat [3]. To‘plamda 16 838 ta anomaliyaga ega 11 175 629 log xabarlari mavjud.



1-rasm. Anomaliyalarni aniqlashda K-means algoritmi

Log parsing – loglarni qayta ishlash va formatlash

PoCni amalga oshirish uchun parsing asoslangan tahlil qilish usuli qo‘llaniladi. IPLoM parserining hech qanday parametrlarini o‘zgartirmasdan ma’lumotlar to‘plamini tahlil qilishni amalga oshirish, ya’ni, standart qiymatlardan foydalanish 156 ta namunani hosil qiladi [4].

Xususiyatlarni ajratib olish

Har bir jurnal xabarida blok identifikatori (block_id) mavjud. Ushbu block_id larning har biri uchun tegishli block_id bilan har bir hodisa sodir bo‘lgan sonini o‘z ichiga olgan hodisa vektori yaratilgan.

Qayta ishlash

Anomaliyalarni aniqlashdan oldin hodisa matritsasi har xil turdagi qayta ishlov berishdan o‘tadi.

Xususiyatlarni standartlashtirish

Hodisa vektorlaridagi xususiyatlar odatda taqsimlanmagan. Ularni normal taqsimlash Scikit-learn’s StandardScaler yordamida standartlashtiriladi.

Xususiyatlarni normallashtirish

Xususiyatlar Scikit-learn kutubxonasi yordamida birlik shaklga aylantiriladi, ya’ni normallashtiriladi. Ushbu kutubxona yordamida uch xil turdagi normallashtirishni amalga oshirish mumkin: l_1, l_2 va max. l_1-normadan foydalanilganda, normallashtirilgan vektor vektorni barcha elementlarning yig‘indisiga bo‘lish yo‘li bilan hisoblanadi. N kattalikdagi x vektor uchun l_1-norma quyidagicha hisoblanadi:

$$z = \sum_{i=1}^{N} |x_i|$$

l_2-norma vektor elementlari yig‘indisining kvadrat ildizini olish yo‘li bilan hisoblanadi:

$$z = \sqrt{\sum_{i=1}^{N} |x_i^2|}$$

Maksimal norma vektorning eng katta elementini hisobga olgan holda hisoblanadi:

$$z = max\ (x_i)$$

Keyin x vektori z ga bo‘linadi va normallashtirilgan vektor olinadi.

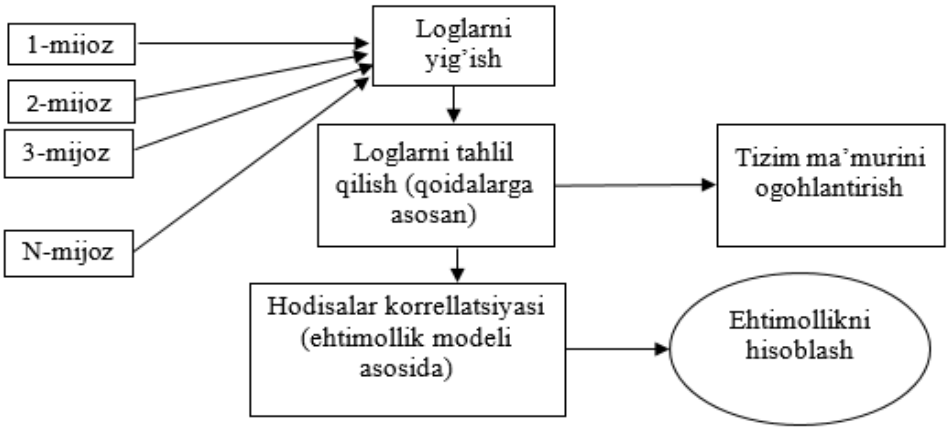
Asosiy komponentlarni tahlil qilish (PCA). PCA hodisa vektor o‘lchovlari sonini kamaytirish va bu natijalarga qanday ta’sir qilishini tekshirish uchun bir nechta testlarda ishlatiladi. PCA dan foydalanishda, o‘lchamlarni kamaytirish jarayonida katta miqdordagi raqamli ma’lumotlar yo‘qolsa yoki siqilsa, aniqlik, xotira va boshqalar bo‘yicha natijalarga salbiy ta’sir ko‘rsatishi mumkin. Boshqa tomondan, bu vaqt murakkabligiga ijobiy ta’sir ko‘rsatishi mumkin.

Anomaliyani aniqlash. Anomaliyalarni aniqlash Scikit-learn dan K-means klasterlash usulidan foydalanadi [4].

Log fayl hodisalarini tahlillashning korrelyatsion yondashuvi

Axborot aktivlarini himoya qilish va tarmoq ishchanligini saqlash uchun standartlashtirilgan log fayli monitori/filtrini ishlab chiqilgan. U to‘rtta moduldan iborat. 2-rasmda taklif qilingan tizimning arxitekturasi ko‘rsatilgan. Har bir modul o‘ziga xos, aniq belgilangan vazifaga ega. Har bir modul ma’lumotlarni oldinga uzatadi. Quyida to‘rtta modul keltirilgan:

- Modul 1: Loglarni yig‘ish
- Modul 2: Loglarni tahlil qilish
- Modul 3: Hodisalar korrellatsiyasi
- Modul 4: Ehtimollikni hisoblash



2-rasm. Tizimning umumiy arxitekturasi

Modul1: Loglarni yig‘ish. Log fayllari ushbu operatsiyada muhim ma’lumotlar manbai hisoblanadi. Ko‘p turli manbalardan ma’lumotlarni to‘plash tizim ma’murlari va xavfsizlik bo‘yicha mutaxassislarga tarmoqning hozirgi holati haqida tasavvur beradi.

Modul 2: Loglarni tahlil qilish. Turli manbalardan juda ko‘p jurnal yozuvlari to‘planadi. Kirish hodisalarini formatidan qat’iy nazar boshqarish uchun ularni tanib olish uchun muntazam ifoda ishlatiladi.

Modul 3: Hodisalar korrellatsiyasi. Hisoblash muhitida hodisa atamasi nima va qachon sodir bo‘lganligini bildiruvchi xabarga nisbatan qo‘llaniladi. Shuning uchun, bitta tizim yoki tizimlar guruhidagi hodisa yoki hodisalar ketma-ketligini aniqlash insayderlarni aniqlash uchun foydalidir.

Modul 4: Ehtimolliklarni hisoblash. Samaradorlik hujumni aniqlash tizimining xususiyatlaridan biridir. Bu tizim qanday darajada hujumlarni aniqlay olishini va noto‘g‘ri signallarni qanchalik yaxshi yo‘q qilishni aniqlaydi. Hujumlarni qidirish uchun ehtimollik usuli qo‘llaniladi. S.Axelsson hodisa chastotasining ahamiyatini ehtimollik va yolg‘ondan tasdiqlash ko‘rsatgichini bilan hisoblashni taklif etgan. A hodisasining ehtimoli A uchun hisoblangan natijalar

sonining A uchun barcha kiruvchi qiymatlar soniga nisbati sifatida aniqlanadi [5]:

$$P(A) = \frac{A\ uchun\ ijobiy\ natijalar\ soni}{Mumkin\ bo'lgan\ natijalarning\ umumiy\ soni}$$

Log jurnalida mavjud bo‘lmagan hodisalarning ko‘pligi (xavfsiz faoliyat) tufayli faqat bir nechta hodisalar (buzilishlar) qiziqish uyg‘otadi. Bayes teoremasi shartli ehtimollikning qo‘llanilishini ifodalaydi.

Log fayllarni imzoga asoslangan aniqlash

Muayyan zararli dastur fayli yoki fayllar oilasi uchun imzo yaratish uchun xavfsizlik bo‘yicha tahlilchiga ishlash uchun faylning bir yoki bir nechta namunalari kerak bo‘ladi. Bunday namunalar asosan zararlangan kompyuterlardan, darknetdan va zararli dastur mualliflari o‘z ishlarini sotadigan boshqa joylardan yoki xavfsizlik tadqiqotchilari ma’lum zararli dasturlar fayllarini almashishi mumkin bo‘lgan umumiy zararli dasturlar omborlaridan to‘planishi mumkin. Xavfsizlik mutaxassislari uchun mavjud bo‘lgan ba’zi mashhur zararli dasturlar omborlari orasida VirusTotal, Malpedia va MalShare mavjud.

SDN tarmoqlarida log fayllarni tahlil qilish usullari

Log tahlili - bu jurnallar deb ataladigan kompyuter tomonidan yaratilgan hodisalar yozuvlarini ko‘rish, sharhlash va tushunish jarayoni. Jurnallar bir qator dasturlashtiriladigan texnologiyalar, jumladan, tarmoq qurilmalari, operatsion tizimlar, ilovalar va boshqalar tomonidan yaratilgan. Jurnal tizimda sodir bo‘layotgan harakatlarni tavsiflovchi vaqt ketma-ketligidagi bir qator xabarlardan iborat. Jurnal fayllari faol tarmoq orqali jurnal kollektoriga uzatilishi yoki keyinchalik ko‘rish uchun fayllarda saqlanishi mumkin.

Log tahlilini amalga oshirish bosqichlari. Loglar ilovalar to‘plami va infratuzilmaning samaradorligi va holatini ko‘rish imkonini beradi, bu esa ishlab chiqish guruhlari va tizim ma’murlariga muammolarni tashxislash va tuzatishni osonlashtiradi. Jurnalni tahlil qilish asosiy besh bosqichli jurnalni boshqarish jarayoni sifatda ifodalanadi va bular:

1. vositalar va yig‘ish - tizimning istalgan qismidan ma’lumotlarni yig‘ish uchun kollektor sozlanadi.
2. markazlashtirish va indekslash - qidiruv va tahlil jarayonini soddalashtirish uchun barcha jurnal manbalaridan ma’lumotlarni markazlashtirilgan platformaga integratsiyalash.
3. qidiruv va tahlil - namunani aniqlash, normallashtirish, teglash va korrelyatsiya tahlili kabi tahlil usullari qo‘lda yoki o‘rnatilgan dasturiy vosita yordamida amalga oshirilishi mumkin.
4. monitoring va ogohlantirish – Machine learning va tahlil qilish orqali AT tashkilotlari real vaqt rejimida avtomatik jurnal monitoringini amalga oshirishi mumkin, bu esa muayyan shartlar bajarilganda ogohlantirishlarni yaratadi.
5. hisobotlar va uskunalar paneli - Optimallashtirilgan hisobotlar va uskunalar paneli jurnalni tahlil qilish dasturining asosiy xususiyatlari hisoblanadi [6].

Loglarni tahlil qilish funksiyalari va usullari. Jurnalni tahlil qilish xususiyatlari foydalanuvchilarga jurnallardan ma’lumotlarni tartibga solish va chiqarishga yordam berish uchun ma’lumotlarni manipulyatsiya qiladi. Quyida jurnalni tahlil qilishning eng keng tarqalgan usullaridan bir nechtasi keltirilgan.

- Normalizatsiya - xabarning qismlari bir xil formatga o‘zgartiriladigan ma’lumotlarni boshqarish usuli. Jurnal ma’lumotlarini markazlashtirish va indekslash jarayoni normallashtirish bosqichini o‘z ichiga oladi.
- Namunalarni tanish – Machine learning ilovalari endi kiruvchi xabarlarni namunalar kitobiga solishtirish va “anomal” va “normal” jurnal xabarlarini farqlash uchun jurnalni tahlil qilish dasturida amalga oshirilishi mumkin. Bunday tizim oddiy jurnal yozuvlariga ahamiyat qaratmaydi, ammo anomal yozuv topilganda ogohlantirish yuboradi.
- Tasniflash va teglash - Jurnal tahlilining bir qismi sifatida bir xil turdagi jurnal yozuvlarini guruhlaydi.

- Korrelyatsiya tahlili - voqea sodir bo‘lganda, u turli manbalardan olingan jurnallarda aks etishi mumkin. Korrelyatsiya tahlili - bu turli tizimlardan jurnal ma’lumotlarini yig‘ish va ma’lum voqea bilan bog‘liq bo‘lgan har bir alohida tizimdan jurnal yozuvlarini aniqlashning analitik jarayoni.

Kiberxavfsizlikda loglar tahlili. Kiberxavfsizlik imkoniyatlarini kengaytirmoqchi bo‘lgan tashkilotlar kibertahdidlarni proaktiv ravishda aniqlash va ularga javob berishga yordam beradigan jurnallarni tahlil qilish imkoniyatlarini rivojlantirishlari kerak. Log tahlili orqali kiberxavfsizlikni samarali nazorat qiladigan tashkilotlar o‘zlarining tarmoq aktivlariga hujum qilishni qiyinlashtirishi mumkin. Kiberxavfsizlik monitoringi, shuningdek, kiberhujumlarning chastotasi va jiddiyligini kamaytirishi, tahdidlarga erta javob berishni rag‘batlantirishi va tashkilotlarga kiberxavfsizlik talablariga javob berishiga yordam berishi mumkin, jumladan:

- ISO/IEC 27002:2013 Axborot texnologiyalari. Xavfsizlik texnologiyalari. Axborot xavfsizligini nazorat qilish amaliyot kodeksi.
- PCI DSS V3.1 (10 va 11 qismlar).

Tashkilot kuzatishi kerak bo‘lgan jurnallar turlarini aniqlash uchun ushbu ro‘yxatni boshlang‘ich nuqtasi sifatida foydalanadi.

3-jadval.

SDN tarmoqlarda kuzatiluvchi asosiy loglar

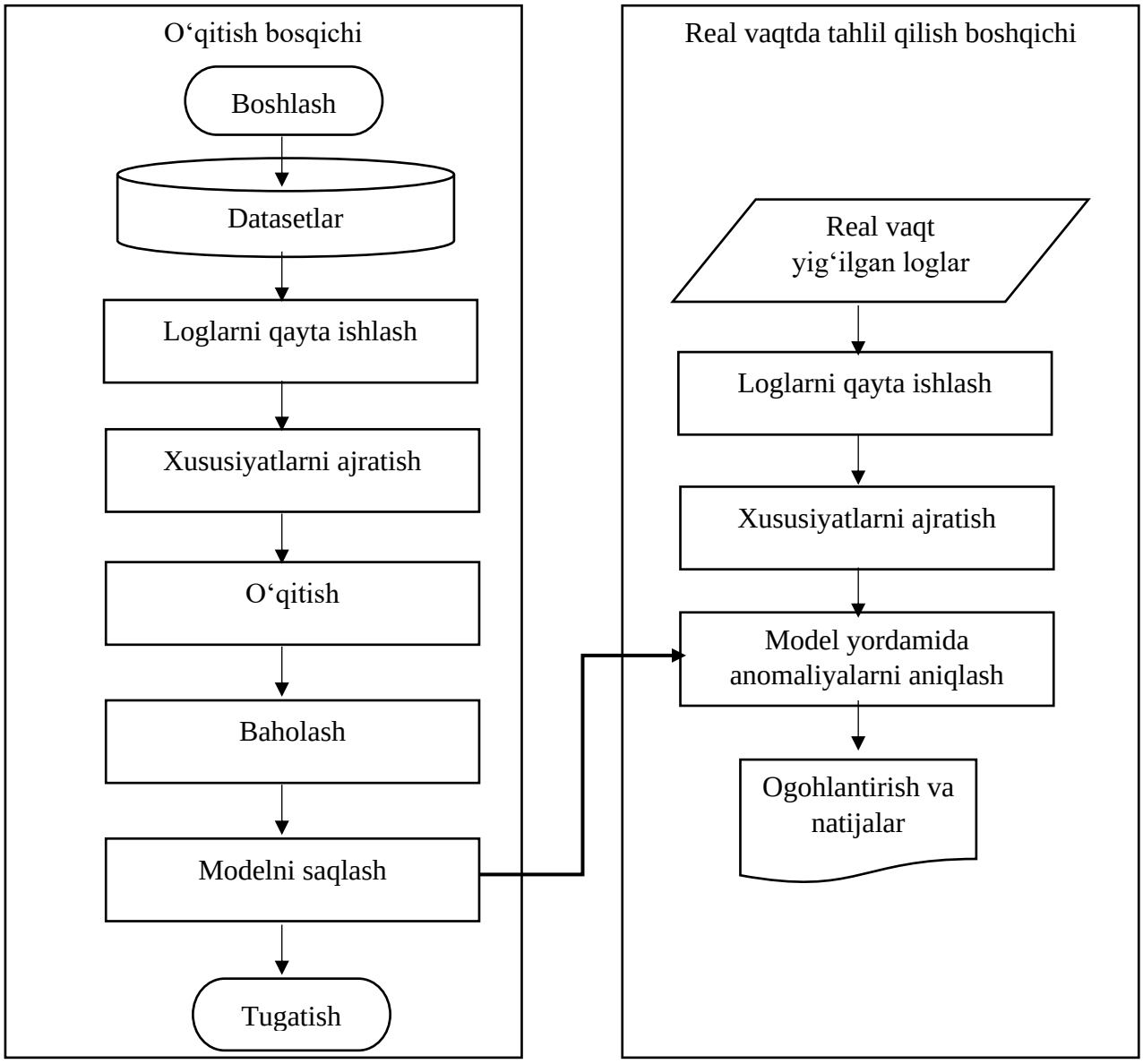
Tizim loglari	Tizim faoliyati loglari. Oxirgi nuqta loglari. Autentifikatsiya loglari. Fizik xavfsizlik loglari.
Tarmoq loglari	Email loglari. Firewall loglari. VPN loglari. Netflow loglari
Texnik loglari	HTTP proxy loglari. DNS, DHCP va FTP loglari. Appflow loglari Web va SQL server loglari.
Kiberxavfsizlik monitoringi loglari	Zararkunanda dasturlardan himoyalash vositalari loglari. Network intrusion detection system (NIDS) loglari. Network intrusion prevention system (NIPS) loglari. Data loss protection (DLP) loglari.

Bularning barchasi uchun Hodisalar jurnali katta hajmdagi ma’lumotlarni yaratishi mumkin va jurnallarni samarali qayta ishlash uchun katta xarajatlar va resurslarni talab qilinadi. Mutaxassislar doimiy monitoring uchun eng muhim loglarni aniqlashlari va vaqt va resurslarni tejash uchun avtomatlashtirilgan yoki dasturiy jurnallarni tahlil qilish usullaridan foydalanishadi.

Dasturiy aniqlangan tarmoqlarda loglarni sun’iy intellektli tahlillash orqali anomaliyalarni aniqlash usuli

Dasturning umumiy strukturasi

3-rasmda Dasturiy aniqlangan tarmoqlarda loglarni sun’iy intellektli tahlillash orqali anomaliyalarni aniqlash usulining umumiy arxitekturasi ko‘rsatilgan. Anomaliyalarni aniqlash tizimi to‘rt bosqichni o‘z ichiga oladi: loglarni yig‘ish, logni tahlil qilish, xususiyatlarni ajratib olish va anomaliyalarni aniqlash.



3-rasm. Loglarni sun’iy intellektli tahlil qilish usulining umumiy arxitekturasi

Loglarni yig‘ish. Katta miqyosli tizimlar muntazam ravishda tizim holati va ish vaqti ma’lumotlarini yozib olish uchun loglarni yaratadi, ularning har biri vaqt tamg‘asi va nima sodir bo‘lganligini ko‘rsatadigan log xabarini o‘z ichiga oladi. Ushbu qimmatli ma’lumotlardan turli maqsadlarda foydalanish mumkin (masalan, anomaliyalarni aniqlash), shuning uchun loglar birinchi navbatda keyingi foydalanish uchun yig‘iladi.

Loglarni tahlil qilish. Loglar strukturalanmagan va erkin shakldagi matnni o‘z ichiga oladi. Logni tahlil qilishdan maqsad, yangi jurnallar tuzilishi mumkin bo‘lgan hodisa namunalari guruhini ajratib olishdir. Aniqroq qilib aytganda, har bir log xabari ba’zi bir o‘ziga xos parametrlar (o‘zgaruvchan qism) bilan voqea namunalariga (doimiy qism) ajratilishi mumkin.

Xususiyatlarni ajratib olish. Loglarni alohida hodisalarga ajratgandan so‘ng, machine learning modellarini qo‘llash uchun ularni raqamli xususiyat vektorlariga kodlash kerak. Loglarning har bir ketma-ketligi uchun xususiyat vektorini (hodisalar soni vektori) hosil qilinadi, bu har bir hodisaning sodir bo‘lish soni hisoblanadi. Barcha xususiyat vektorlari birgalikda xususiyat matritsasi, ya’ni hodisalar soni matritsasi hosil qilishi mumkin.

Anomaliyalarni aniqlash. Nihoyat, xususiyat matritsasi o‘qitish uchun machine learning modellariga berilishi mumkin va shu bilan anomaliyalarni aniqlash uchun model yaratiladi. Tuzilgan model yangi kiruvchi loglar ketma-ketligi anomaliya yoki yo‘qligini aniqlash uchun ishlatilishi mumkin [4].

Natijalar

Jurnal ma’lumotlar to‘plami. Ommaviy ishlab chiqarish loglari juda kam ma’lumotlarga ega, chunki kompaniyalar maxfiylik muammolari tufayli ularni kamdan-kam chiqaradi. HDFS va BGL ma’lumotlar to‘plami ishlab chiqarish tizimlaridan to‘plangan va manba domenidagi mutaxassislar tomonidan qo‘lda belgilangan 15,923,592 jurnal xabarlari va 365,298 anomaliya namunalarini o‘z ichiga oladi. Shunday qilib, ushbu teglarni (anomaliya yoki normal) aniqlikni baholash uchun haqiqiy ma’lumot sifatida qabul qilinadi.

HDFS ma’lumotlari Amazon EC2 platformasidan to‘plangan 11,175,629 jurnal xabarlarini o‘z ichiga oladi. HDFS jurnallari ajratish, yozish, replikatsiya, o‘chirish kabi har bir blok operatsiyasi uchun noyob blok identifikatorini yozadi. Shunday qilib, log operatsiyalari session window tomonidan tabiiy tasvirga olinishi mumkin, chunki har bir noyob blok identifikatori loglarni loglar ketma-ketligi to‘plamiga bo‘lish uchun ishlatilishi mumkin. Keyin ushbu loglar ketma-ketligidan xususiyat vektorlarini chiqaradi va 575061 hodisalarni hisoblash vektorlarini yaratadi. Ular orasida 16 838 ta namuna anomaliya sifatida belgilangan.

Testlashda HDFS ma’lumotlar bazasining 100000 log yozuvlariga ega HDFS_100k.log_struktured.csv strukturalangan va anomaliya va normal belgilariga ko‘ra tasniflangan anomaly_label.csv ma’lumotlar to‘plamlaridan foydalaniydi.

Testlash Intel Core i3-9100. 4x3.6 GHz, 8 GB DDR4, Nvidia GeForce GTX 1050, 1 TB HDD, 128 GB SSD, 64 bit Ubuntu 18.04 xususiyatlarga ega amalga kompyuterda oshirilgan.

Anomaliyalarni aniqlash usullarining to‘g‘riligini baholash uchun eng ko‘p qo‘llaniladigan ko‘rsatkichlar bo‘lgan aniqlik, eslab qolish va F1-o‘lchovidan foydalaniladi, chunki ikkala ma’lumotlar to‘plami uchun asosiy aniqlanuvchi qiymat (anomaliya yoki yo‘q) mavjud. Quyida ko‘rsatilganidek, aniqlik to‘g‘ri aniqlangan anomaliyalar foizini o‘lchaydi, eslab qolish aniqlangan haqiqiy anomaliyalar foizini o‘lchaydi va F1 aniqlik va eslab qolishning garmonik o‘rtacha qiymatini o‘lchaydi [1].

$$Aniqlik = \frac{aniqlangan\ anomaliyalar\ soni}{xabar\ berilgan\ anomaliyalar\ soni}$$

$$To‘g‘rilik = \frac{aniqlangan\ anomaliyalar\ soni}{barcha\ anomaliyalar\ soni}$$

$$F1 = \frac{2 * aniqlik * to‘g‘rilik}{aniqlik + to‘g‘rilik}$$

Supervised learning barcha uchta usuli uchun ma’lumotlarning dastlabki 80 foizini o‘quv ma’lumotlari sifatida, qolgan 20 foizini esa test ma’lumotlari sifatida tanlanadi, chunki faqat oldingi voqealar keyingi anomaliyaga olib kelishi mumkin.

4-jadval.

Dastur yordamida anomaliyalarni aniqlash uchun testlash natijalari

Algoritm	Aniqlik	To‘g‘rilik	F1
PCA	0.966	0.363	0.528
Invariant Mining	0.973	0.559	0.710
Decision Tree	0.985	0.427	0.596
Isolation Forest	0.985	0.427	0.596
Clustering	0.967	0.561	0.710
Logistic Regression	0.986	0.433	0.602
SVM	0.986	0.433	0.602

Xulosa

Ushbu maqolada dasturiy aniqlangan tarmoqlarda (SDN) log fayllarini sun’iy intellekt asosida tahlil qilish orqali anomaliyalarni aniqlash

usuli taklif etilgan. SDN tarmoqlarda sodir bo‘layotgan anomaliyalar, jumladan, noqonuniy kirish urinishlari, trafik manipulyatsiyasi, konfiguratsiya xatoliklari kabi hodisalarni o‘z vaqtida aniqlash zarur. Log fayllari tarmoq faoliyatining barcha jihatlari bo‘yicha qimmatli ma’lumot manbai bo‘lib, ularni tahlil qilish kiberxavfsizlik monitoringi uchun asosiy vosita hisoblanadi.

Loglarni tahlil qilishda sun’iy intellekt va mashina o‘qishi algoritmlaridan foydalanish tizimning o‘ziga xos normal holatini o‘rganish va undan chetlanishlarni (anomaliyalarni) aniqlash imkonini beradi. Umuman olganda, loglarni sun’iy intellektli tahlil qilish orqali anomaliyalarni aniqlash SDN tarmoqlarida xavfsizlikni kuchaytirish, hujumlarga erta javob berish va tarmoq faoliyatini doimiy nazorat qilish uchun samarali va kelajakdagi rivojlanish istiqboli yuqori usuldir. Keyingi tadqiqotlarda chuqur o‘rganish (deep learning) modellari, ayniqsa, LSTM va Transformer tizimlari orqali real vaqtda loglarni tahlil qilish samaradorligi yanada oshirilishi mumkin.

Foydalanilgan adabiyotlar ro‘yxati

1. Ming Zhong, Yajin Zhou, Gang Chen, “A Security Log Analysis Scheme Using Deep Learning Algorithm for IDSs in Social Network”, Security and Communication Networks, vol. 2021, Article ID 5542543, 13 pages, 2021. https://doi.org/10.1155/2021/5542543
2. Ricardo Ávila, Raphaël Khoury, Richard Khoury, Fábio Petrillo, “Use of Security Logs for Data Leak Detection: A Systematic Literature Review”, Security and Communication Networks, vol. 2021, Article ID 6615899, 29 pages, 2021. https://doi.org/10.1155/2021/6615899
3. Ambre, A., & Shekokar, N. (2015). Insider Threat Detection Using Log Analysis and Event Correlation. Procedia Computer Science, 45, 436–445. doi:10.1016/j.procs.2015.03.175
4. Mans F. Franzén, Nils Tyrén. Anomaly detection for automated security log analysis – Comparison of existing techniques and tools. Linköping University Department of Computer and Information Science Master’s thesis, 30 ECTS Datateknik 2021 | LIU-IDA/LITH-EX-A--2021/033—SE.
5. Leemans, M., van der Aalst, W. M. P., & van den Brand, M. G. J. (2018). The Statechart Workbench: Enabling scalable software event log analysis using process mining. 2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER). doi:10.1109/saner.2018.8330248.
6. Leemans, M., van der Aalst, W. M. P., & van den Brand, M. G. J. (2018). Recursion aware modeling and discovery for hierarchical software event log analysis. 2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER). doi:10.1109/saner.2018.8330208.
7. A.Makanju, A.N. Zincir-Heywood and E.E. Milios, “Investigating event log analysis with minimum apriori information,” 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), 2013, pp. 962-968.
8. Ghosh, S., et al. (2017). “Anomaly Detection in Software-Defined Networking,” Journal of Network and Computer Applications, vol. 79, pp. 77-85.