Kauno technologijos universitetas

Informatikos fakultetas

Mašininio mokymo metodų taikymas atakoms aptikti naudojant *Cisco Netflow* tinklo įrašus

Baigiamasis magistro studijų projektas

Giedrius Aleksandravičius

Projekto autorius

doc. Rimantas Kavaliūnas

Vadovas

Kaunas, 2021

Kauno technologijos universitetas

Informatikos fakultetas

Mašininio mokymo metodų taikymas atakoms aptikti naudojant *Cisco Netflow* tinklo įrašus

Baigiamasis magistro studijų projektas

Informacijos ir informacinių technologijų sauga (6211BX008)

Giedrius Aleksandravičius

Projekto autorius

doc. Rimantas Kavaliūnas

Vadovas

dr. Šarūnas Grigaliūnas

Recenzentas

Kaunas, 2021

Kauno technologijos universitetas

Informatikos fakultetas

Giedrius Aleksandravičius

Mašininio mokymo metodų taikymas atakoms aptikti naudojant *Cisco Netflow* tinklo įrašus

Akademinio sąžiningumo deklaracija

Patvirtinu, kad:

1. baigiamąjį projektą parengiau savarankiškai ir sąžiningai, nepažeisdama(s) kitų asmenų autoriaus ar kitų teisių, laikydamasi(s) Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo nuostatų, Kauno technologijos universiteto (toliau – Universitetas) intelektinės nuosavybės valdymo ir perdavimo nuostatų bei Universiteto akademinės etikos kodekse nustatytų etikos reikalavimų;

2. baigiamajame projekte visi pateikti duomenys ir tyrimų rezultatai yra teisingi ir gauti teisėtai, nei viena šio projekto dalis nėra plagijuota nuo jokių spausdintinių ar elektroninių šaltinių, visos baigiamojo projekto tekste pateiktos citatos ir nuorodos yra nurodytos literatūros sąraše;

3. įstatymų nenumatytų piniginių sumų už baigiamąjį projektą ar jo dalis niekam nesu mokėjęs (-usi);

4. suprantu, kad išaiškėjus nesąžiningumo ar kitų asmenų teisių pažeidimo faktui, man bus taikomos akademinės nuobaudos pagal Universitete galiojančią tvarką ir būsiu pašalinta(s) iš Universiteto, o baigiamasis projektas gali būti pateiktas Akademinės etikos ir procedūrų kontrolieriaus tarnybai nagrinėjant galimą akademinės etikos pažeidimą.

Giedrius Aleksandravičius

Patvirtinta elektroniniu būdu

Aleksandravičius, G. „Mašininio mokymo metodų taikymas atakoms aptikti naudojant *Cisco Netflow* tinklo įrašus“. Magistro baigiamasis projektas / vadovas doc. Rimantas Kavaliūnas; Kauno technologijos universitetas, Informatikos fakultetas, Kompiuterių katedra.

Mokslo kryptis ir sritis: Kompiuterių mokslai, Informacijos ir informacinių technologijų sauga

Reikšminiai žodžiai: „Cisco Netflow“ tinklas, gilaus mokymosi metodas, klasifikatoriai, sudėtinis klasifikatorius, anomalijų aptikimas, atakų aptikimas.

Kaunas, 2021. 64 p.

SANTRAUKA

Šiais laikais labai svarbi tinklo saugos problema. Kartais visai to nežinodami patiriame bandymus įsilaužti į vidinius įmonės tinklus, ko pasekoje yra galimybė patirti nuostolius dėl sistemos neveiklumo ar pačio kompiuterio užvaldymo. Šioms problemoms užkirsti, naudojamos anomalijų aptikimo sistemos, kurios aptinka šias atakas ir praneša administratoriui, kuris imasi tolimesnių veiksmų.

Darbe ištirta 9 metodai anomalijoms aptikti. Analizėje pastebėta jog nėra tokio metodo, kuris naudotų mažą mokymosi duomenų kiekį. Pasinaudojus darbe aprašyta projektavimo faze, buvo nustatyti reikalavimai klasifikatoriams, sudaryta norima architektūra bei pradėtas realizacijos etapas, kuriame buvo realizuotas gilaus mokymosi metodas kartu su sudėtiniu klasifikatoriumi, kurie naudoja sintetinį duomenų rinkinį. Šie metodai buvo palyginti pagal darbe nurodytus vertinimo kriterijus.

Aleksandravičius, Giedrius. Machine Learning Based Attack Detection Using *Cisco Netflow* Network Records: Master's thesis / supervisor doc. Rimantas Kavaliūnas. The Faculty of Informatics, Kaunas University of Technology.

Research area and field: Computer science, Information and Information Technology Security

Key words: "Cisco Netflow" network, deep learning algorithm, classification, stacked classifier, anomaly detection, attack detection.

Kaunas, 2021. 64 p.

SUMMARY

Network security is a very important issue these days. Sometimes, unknowingly, we experience attempts to break into the company's internal networks, as a result of which there is a possibility of incurring losses due to system inactivity or taking control of the computer itself. To prevent these problems, anomaly detection systems are used that detect these attacks and notify the administrator, who takes further action.

The paper analyzes 9 methods for detecting anomalies. The analysis found that there is no method that uses a small amount of learning data. Using the design phase described in the work, the requirements for classifiers were determined, the desired architecture was created and the implementation phase was started, in which the deep learning method was implemented, together with a composite classifier. These methods were compared according to the evaluation criteria given in the work.

TURINYS

LENTELIŲ SĄRAŠAS

PAVEIKSLŲ SĄRAŠAS

TERMINŲ IR SANTRUMPŲ ŽODYNAS

NADS (angl. *Network Anomalies Detection system*) – tinklo anomalijų aptikimo sistema.
NIDS (angl. *Netvork Intrusion Detection system*) – tinklo įsilaužimų aptikimo sistema.
IDS (angl. *Intrusion Detection system*) – įsilaužimų aptikimo sistema.
DoS (angl. *Denial of service*) – paslaugos trikdymo ataka.
DDoS (angl. *Distributed denial of service*) – paskirstyta paslaugos trikdymo ataka.
ANN (angl. *Artificial Neural Network*) – dirbtinis neuroninis tinklas.
KNN (angl. *K-nearest neighbor)* – K-artimiausio kaimyno metodas.
SGD (angl. *Stochastic Gradient Descent*) – stochastinio gradiento nusileidimas.
MLP (angl. *Multilayer Perceptron)* – daugiasluoksnis perceptronas.
SVM (angl. *Support Vector Machine)* – pagalbinio vektoriaus mašina.

ĮVADAS

Informacijos ir informacinių technologijų sistemų saugos studento Giedriaus Aleksandravičiaus darbas. Šis darbas priklauso Informacijos ir informacinių technologijų saugos krypties eksperto 6211BX008 studijų programai.

Darbo problematika ir aktualumas

Dauguma įmonių remiasi senomis IT sistemomis, kurias sudaro perimetro apsauga ir galinių taškų apsauga. Pasaulyje, kuriame grėsmės turi daugiau galimybių nei bet kada anksčiau apeiti tradicinius sprendimus ir pasislėpti, kur 70% atakų kyla iš vidinio tinklo, šio požiūrio nebepakanka. Todėl reikia apsaugoti savo sistemas ir duomenis nuo sudėtingų, nuolat kintančių atakų, kurių tradiciniai sprendimai neaptinka.

NBAD sprendimai nuolat stebi tinklo srautą, analizuodami ryšį, norėdami ieškoti anomalijų ir atskleisti įtartiną elgesį. Tai įgalina reaguoti į dar nežinomas saugumo grėsmes, kurių negalima aptikti kitomis technologijomis. Daugelis įmonių pradėjo pripažinti faktą, kad NBAD sistemos yra neatsiejamos tinklo saugos užtikrinimo dalis.

Tinklo elgesio anomalijų aptikimo sistemos ir tinklo elgsenos analizė vis dažniau pasitaiko kibernetinio saugumo architektūrose, o vis daugiau tinklo administratorių siekia šių sistemų, kad padidintų savo nuosavybės saugumą ir apsaugos lygį. Kai dirbtinio intelekto technologijos, tokios kaip mašininis mokymasis, neuroniniai tinklai ir predikatyvūs algoritmai, pastaraisiais metais mato daug daugiau tyrimų, plėtros ir finansavimo, NBAD sistemų galia ir intelektas per ateinančius kelerius metus gali išaugti.

Darbo tikslas ir uždaviniai

Darbo tikslas yra ištirti anomalijų aptikimu paremtų tinklo atakų aptikimo metodų efektyvumą ir suprojektuoti tinkamiausią metodą konkretaus tipo atakoms atpažinti iš „Cisco Netflow“ įrašų gaunamų savybių. Šiam tikslui yra iškelti šie uždaviniai:

- atlikti analizę ir ištirti esamus tinklo elgsenos anomalijų aptikimo metodus;
- suprojektuoti tokiam tyrimui veiksmingą modelį, kuriame galima tirti atskirus metodus ir jų kombinacijas, kaitalioti apsimokymo duomenų rinkinio savybes;
- realizuoti suprojektuotą tyrimo modelį programiškai;
- atlikti eksperimentus su pasirinktais metodais 4 atakų tipams ir palyginti gautus rezultatus.

Darbo struktūra

Numatoma darbo struktūra turės skyrius tokius skyrius: analizė, projektavimas, realizacija, tyrimas ir darbo išvados. Toliau trumpai aprašoma kiekvienas šio darbo skyrius:

Analizės skyriuje aprašoma problematika, galimi tinklo anomalijų aptikimo sistemos metodai, pateikiama informacija apie panašias sistemas.

Projektavimo skyriuje aprašomas anomalijų aptikimo metodo pasiūlymas.

Realizacijos skyriuje pasiūlytas metodas įgyvendinamas.

Tyrimo skyriuje aprašomi metodų įvertinimai ir palyginimas su kitais metodais.

1 PROBLEMINĖS SRITIES ANALIZĖ

1.1 Analizės tikslas

Analizės tikslas – ištirti anomalijos tinkle galimus elgesius ir pasireiškimo būdus, atlikti lyginamąją tinklo anomalijų aptikimo sistemų analizę, įvardinti galimas problemas.

1.2 Tyrimo objektas, sritis ir problema

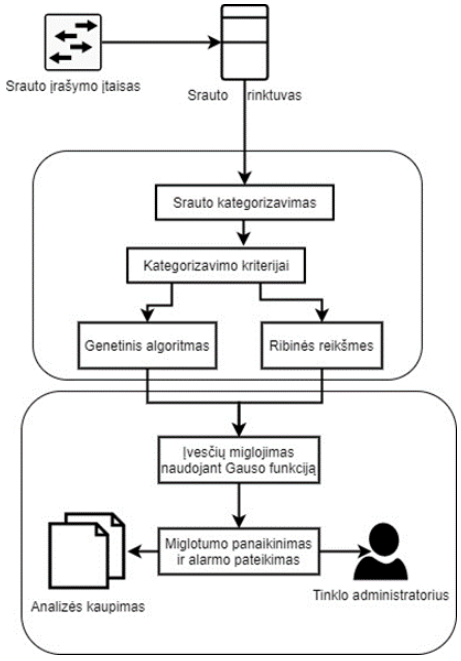
Viena iš tinklo anomalijų aptikimo sistemų problemų yra jų kalibravimo bei apmokymo laiko kiekis, nes sistemos apmokymas užtrunka didelį laiko tarpą, jie naudojama daug duomenų apmokymui. Antra, daugelis įmonių vis dar naudoja rankiniu būdu konfigūruojamas anomalijų aptikimo sistemas su statiniais slenksčiais, o jų reikšmės yra parenkamos administratorių nuožiūra.

1.2.1 Situacija Lietuvoje ir pasaulyje

Šiais laikais vis dar aktualu tirti tinklo anomalijas, nes vis besiplečiantys interneto tinklai ir atsirandančios naujos technologijos, didina tikimybę, kad juose gali būti vykdomos tinklo atakos, kurios yra anomalijos ar jos požymis.

LITNET infrastruktūroje (magistralėje, techniniuose centruose ir jungimo mazguose) įrengta daugiau kaip 600 nuotoliniu būdu valdomų ir stebimų įrenginių. Pasaulio mastu ši infrastruktūra yra milžiniška, į ją kasmet investuojama milijonai lėšų, kurios naudojamos tinklo saugos sistemų plėtimui. Kadangi kiekviena tinklo paslaugas teikianti kompanija nori žinoti, kas vyksta jų stebimuose tinkluose, todėl visos šios įmonės turi nuosavas tinklo elgsenos anomalijų aptikimo sistemas (sutrumpintai NBAD).

NBAD puikiai atvaizduoja šaltinyje [1] pavaizduotas sistemos veikimo principas ir sudėtinių dalių schema (žr. 1.1 pav.). Šioje schemoje atvaizduotas pagrindinis anomalijų aptikimo principas, kuris sudarytas iš tinklo srauto kaupiklio, kuris atrenka tinklo srauto duomenis ir juos perduoda į tinklo srauto analizatorių, kuris pagal metodo apmokymo duomenis



1.1 pav. Anomalijų aptikimu pagrįsta įsibrovimo aptikimo sistema, apjungianti genetinį algoritmą ir miglotąją logiką [1]

1.3 Tinklo elgsenos anomalija ir normalus veikimas

Tinklo anomalijos paprastai pasireiškia, kai tinklo elgsena skiriasi nuo įprasto tinklo elgesio. Anomalijos gali atsirasti dėl įvairių priežasčių, tokių kaip netinkamas tinklo įrenginių veikimas, netinkama tinklo paslaugų ir operacinių sistemų konfigūracija, tinklo perkrova, paslaugos trikdymo atakos, netinkamai įrengtos vartotojų programinės įrangos, vartotojų pastangos prisijungti prie tinklo ir surinkti informaciją, apie jo įrenginius bei tinklo įsibrovimus, kurie trikdo įprastą tinklo paslaugų teikimą. Šie anomalūs įvykiai sutrikdys normalų, kai kurių išmatuojamų tinklo duomenų, elgesį. Šis apibrėžimas paminėtas šaltinyje [2].

Norint nustatyti, kas yra normalus veikimas tinkle, reikia ilgai ir atidžiai stebėti tinklą. Įprasto tinklo elgesio apibrėžimas išmatuotiems tinklo duomenims priklauso nuo kelių tinklo veiksnių, tokių kaip tiriamo tinklo dinamika, atsižvelgiant į srauto apimtį, turimų tinklo duomenų tipą ir tinkle veikiančių programų tipus. Tikslus normalios tinklo elgsenos modeliavimas vis dar yra aktyvi tyrimų sritis, ypač tinklo srauto modeliavimas.

Kai kurie įsibrovimai ir neturi reikšmingos įtakos tinklo srautui. Kiti išpuolių tipai yra pagrįsti didelio skaičiaus paketų, kurie išsiskiria savo srautu lyginant su normaliu veikimu, perdavimu, kaip DoS atakų atveju. Paprastai anomalija skiriasi nuo didelio paketų skaičiaus, nors didelis paketų skaičius taip pat sukelia tinklo srauto anomalijas. Didelis procentas paketų pablogina tinklo veikimą. Yra ir kitų rūšių atakų, kurios transliuoja srautą aptikdamos tiesioginius pagrindinius kompiuterius tinkle. Tinklo anomalijas taip pat gali sukelti netyčiniai veiksmai. Norint surinkti šiuos duomenis reikalingi tinklo srautų registratoriai, iš kurių populiariausias „Cisco Netflow“.

Todėl išskiriamos 3 pagrindinės anomalijų rūšys, kurios nurodytos straipsnyje [3]:

1. tinklo eksploatavimo anomalijos: tai tinklo įrenginio veikimo sutrikimai, reikšmingi tinklo elgsenos skirtumai, atsirandantys dėl konfigūracijos pakeitimų (pvz., pridedant naują įrangą arba nustatant normos ribas) ir plato elgesį sukeliantį srautą pasiekiantį aplinkos ribas. Šios kategorijos anomalijos vizualiai išsiskiria stačiais, beveik akimirksniu besikeičiančiais bitų perdavimo spartos pokyčiais, po to seka stabilus srautas per tam tikrą laiką, tačiau jau ne toks koks buvo;
2. „flash mob“ anomalijos: mūsų aplinkoje šios kategorijos anomalijos paprastai atsiranda dėl programinės įrangos išleidimo arba dėl išorinio susidomėjimo svetaine dėl tam tikro nacionalinio viešumo. „Flash minios“ elgesys išsiskiria tuo, kad greitai auga tam tikro tipo srautai (pvz., FTP srautai) į žinomą adresą, laikui bėgant srautas palaipsniui mažėja;
3. piktnaudžiavimo tinklu anomalijos: dviejų tipų tinklų piktnaudžiavimas, kurį galima nustatyti naudojant srautus, yra DoS užtvindymo išpuoliai ir prievadų nuskaitymai. Piktnaudžiavimo tinklu anomalijos skiriasi nuo tinklo eksploatavimo ir „Flash minios“ anomalijų tuo, kad jos ne visada lengvai pastebimos matuojant bitų ar paketų spartą. Tačiau srauto skaičiavimai aiškiai rodo piktnaudžiavimo veiklą, turint daug skirtingų šaltinio adresų ir prievadų porų, nes kiekviena jungtis rodoma kaip atskiras srautas.

1.4 Parametrai naudojami tinklo anomalijos ir teisingos elgsenos profilių aprašymui

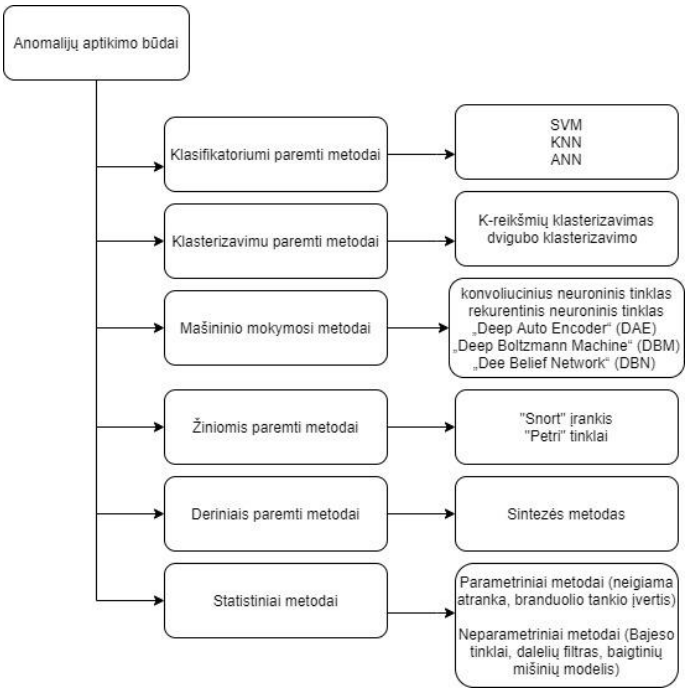
Norint aptikti anomalijas reikia stebėti tokius požymius tinkle, kuriuos stebi komercinės anomalijų aptikimo sistemos kaip „Allot Anomaly Detection“ [4]:

1. didelis paketų kiekis;

2. mažas paketo dydis arba didelis paketo dydis;
3. susibūrimas tinkle (angl. *Fan-in*) (daug IP į vieną IP adresą), tipinė DDoS ataka;
4. išsiplėtimas tinkle (angl. *Fan-out*), (vienas IP į daug IP adresų) DDoS ataka;
5. spiečiai (angl. *swarms*), (daug IP į daug IP adresų) DDoS ataka;
6. DoS (vienas IP į vieną IP);
7. TCP paketais paremti incidentai (SYN, FIN, ACK, RST, netinkamos vėliavėlių kombinacijos);
8. UDP paketais paremti incidentai;
9. ICMP paketais paremti incidentai ;
10. kiti incidentai ;
11. incidentai įtraukiantys padalintus, neteisingai sudarytis paketus.

1.5 Tinklo anomalijų aptikimui naudojami metodai.

Tinklo anomalijų aptikimui galima naudoti daug metodų, pagal metodo mokymosi tipą jie yra skirstomi į šias kategorijas [5]: klasifikatoriumi paremti ir grupavimu paremti metodai, mašininio mokymosi metodai, žiniomis pagrįsti metodai, kombinacijomis paremti metodai ir statistiniai metodai. Šias kategorijas galime pamatyti 1.2 pav.



1.2 pav. Tinklo anomalijų aptikimo būdai ir metodai

1.5.1 Klasifikatoriumi paremti metodai tinklo anomalijoms aptikti

Klasifikatoriumi paremti metodai priskiria duomenis į tam tikras grupes, kurie yra mokymosi duomenyse, pavyzdžiui, turime 2 kategorijas „1“ ir „2“, klasifikatorius duomenis skirsto į šias dvi grupes pagal tam tikrus požymius. Klasifikuodami tinklo srauto elgseną galime skirtyti į dvi klases (t. y., normalią elgseną ar ataką) arba į klasių rinkinį (t. y., kai kiekviena ataka yra klasė). Populiariausi klasifikacija pagrįsti metodai, taikomi NADS, yra palaikymo vektoriaus (SVM), artimiausio kaimyno (KNN), taip pat dirbtinio neuronų tinklo (ANN) metodai.

– Palaikančių vektorių metodas susideda iš dviejų etapų: mokymosi duomenys iš originalios erdvės perkeliamos į aukštesnę erdvę, paremtą branduolio funkcijomis, kurios šiuos duomenis

paverčia į linijiškai atskiriamus iš neatskiriamų. Vienos klasės SVM naudoja tik teisėto tinklo duomenų mokymo rinkinį ir bet kokius nukrypimus nuo įprastų šablonų laiko anomalija. Naudojant vienos klasės SVM metodą, buvo norima aptikti nulinės dienos atakas, nepriklausančias įprastai treniruočių klasei. Tačiau šiai technikai panaudoti dažnai reikėjo ilgai mokyti sistemą dideliam duomenų kiekiui. Panašiai buvo pasiūlyta ir NADS sistema, apimanti hierarchinę klasterizaciją ir SVM, siekiant sumažinti mokymo etapo apdorojimo laiką ir padidinti aptikimo greitį. Vėliau buvo pasiūlytas mažiausio kvadrato SVM metodas, kad būtų galima suprojektuoti lengvą NADS, pasirenkant reikšmingas tinklo duomenų savybes ir aptikti anomalijas.

– KNN pagrįsta NADS sukuria normalų tinklo profilį ir visus nukrypimus nuo jo traktuoja kaip išpuolį. Tai yra galingas aptikimo metodas, nes nereikalauja pritaikyti parametrų mokymo etape. KNN metodas buvo naudojamas kuriant patikimą NADS sistemą, pagrįstą jo galimų funkcijų keistenybėmis ir izoliacijos priemonėmis, kurios galėtų veiksmingai nustatyti tinklo anomalijas. Artimiausio kaimyno metodas dažnai reikalauja daug laiko ir didelių saugyklų, kad būtų galima klasifikuoti spartų tinklo srautą.

Kuriant NADS, taip pat buvo taikomi kiti klasifikavimo būdai, pavyzdžiui, sprendimų medis, regresijos modeliai ir neaiški logika (angl. *Fuzzy Logic*). Tačiau apskritai klasifikacija pagrįstos NADS labai remiasi prielaida, kad kiekvienas klasifikatorius turi būti derinamas atskirai ir visada sunaudoja daugiau išteklių nei statistiniai metodai. Galiausiai, jei šie metodai nesukuria normalių modelių, jie nesugeba aptikti naujų atakų. Svarbu pažymėti, kad dauguma klasifikavimo būdų buvo įvertinti naudojant senus duomenų rinkinius ir jų veikimas bus prastesnis naudojant naujesnius duomenų rinkinius.

Tolesniuose poskyryje yra giliau aptariami mokymosi metodai aprašyti šaltinyje [6].

1.5.1.1 K artimiausio kaimyno algoritmas

K-artimiausio kaimyno algoritmas (KNN) klasifikuoja naujus egzempliorius, egzistuojančius tam tikrame duomenų rinkinyje, pagal artimiausius jų mokymo pavyzdžius ypatybių erdvėje. KNN yra paprastas algoritmas, parodantis patikimą triukšmingų treniruočių duomenų ar didelio duomenų rinkinio patikimumą. Verta paminėti, kad KNN ir jo variantai yra plačiai naudojami aptinkant kenkėjiškas programas, įsilaužimo aptikimui ir šlamšto aptikimui. Paprastai algoritmas suranda artimiausias instancijas, apskaičiuodamas atstumą tarp naujų egzempliorių ir visų mokymo egzempliorių. Pavyzdžiui, tegul X ir Y yra du požymių vektoriai, kurių matmuo n. Euklido atstumas tarp šių dviejų bruožų vektorių yra apibrėžtas žemiau.

$$d_{XY} = \max(|X_i - Y|), i \in n \qquad (1)$$

Naujoji instancija yra klasifikuojama remiantis jos artimiausių instancijų balsų dauguma. Todėl jis bus priskirtas klasei, kurios etiketės yra dažniausiai naudojamos. KNN veikimui įtakos turi atstumo metrika, naudojama algoritmo, kartu pasirenkant optimalią parametro k vertę. Mažas k padidina atskirų atvejų poveikį. Didelis k padidina duomenų rinkinyje pateikto triukšmo patikimumą. K vertė paprastai nustatoma pasitelkiant kryžminį įteisinimą arba pritaikomuosius metodus. Šiame tyrime mes naudojame standartinę šio algoritmo versiją, naudodami Euklido atstumą kaip atstumo metriką ir taikydami kryžminį patvirtinimą, kad nustatytume optimalią parametro k vertę.

1.5.2 Klasterizacija paremti metodai tinklo anomalijoms aptikti

Klasterizacijos metodai yra neprižiūrimi mašininio mokymosi mechanizmai, priskiriantys duomenų taškų grupes pagal panašias šių taškų savybes, tokias kaip atstumas ar tikimybės matavimai. Nors yra skirtingų klasifikavimo metodų, populiariausi NADS sistemose naudojami tipai yra reguliarūs ir dvigubo klasterizavimo (angl. *co-clustering*) strategija su skirtumais tarp jų strategijų apdorojant anomalijos požymius duomenų rinkinyje. Tiksliau tariant, reguliarus klasterizavimas kaip K-reikšmių klasterizavimas (angl. *K-means clustering*) renka duomenų taškus iš duomenų rinkinyje pastebėtų ypatumų, o dvigubas klasterizavimas kartu atsižvelgia ir į stebėjimus ir į ypatybes, taip išskirdamas skirtingus klasterius.

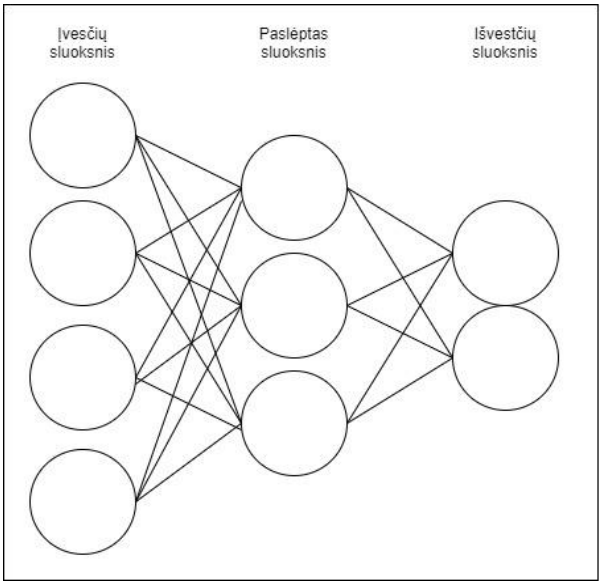
Naudojant klasterius anomalijoms nustatyti paprastai daromos trys pagrindinės prielaidos. Pirma, kadangi teisėti duomenų pavyzdžiai dažnai patenka į klasterį, o atakų duomenys ne. Taikant NADS metodiką, grupavimas identifikuoja visus duomenų atvejus, kurie nepatenka į teisėtą klasterį, kaip išpuolius, o triukšmo duomenys taip pat laikomi anomalija. Šios prielaidos trūkumas yra tas, kad klasterizavimo metodai negali būti optimizuoti, norint nustatyti anomalijas, nes pagrindinis klasterizacijos algoritmo tikslas yra apibrėžti grupes. Antra, teisėti duomenų egzemplioriai paprastai yra arčiausiai klasterio centro reikšmių (angl. *centroid*), o anomalijos dažnai yra toli nuo jų. Taikant šią prielaidą, taškai, esantys toliausiai nuo klasterio centro, laikomi anomalijomis, dauguma jų siūlomi NADS projektavimui, tuo tarpu, jei anomalijos yra įprastuose klasteriuose, jų negalima teisingai nustatyti. Trečioji prielaida, siekiant išspręsti šią problemą, yra tai, kad teisėti duomenų egzemplioriai suskirstomi į didelius ir tankius klasterius, o anomalijos – į mažus. Remiantis šia prielaida, klasteriams priklausantys duomenų stebėjimai identifikuojami su dydžių ir (arba) tankių duomenų stebėjimais pagal bazinę liniją, laikomą anomalijomis.

NADS tinkluose, kuriuose teisingi duomenys buvo grupuojami naudojant K-reikšmių klasterizavimo metodą, o tada kiekvienam klasteriui buvo apskaičiuojamas atskaitos taškas, šie taškai klasifikuojami kaip išpuoliai, jei jie buvo mažesni už tam tikrą ribinę vertę. Taip pat buvo pasiūlytas NADS, skirtas dideliems tinklo duomenų rinkiniams, naudojant medžiais pagrįstais klasterizavimo ir ansambliais pagrįstus metodus, siekiant pagerinti tikslumą realioje tinklo aplinkoje. Buvo išanalizuota ir įvertinta K-reikšmių klasterizavimo hierarchinė ir neaiški C-reikšmių klasterizavimo metodika, naudojama NADS sistemai kurti. Tačiau ši sistema negalėjo efektyviai dirbti su nesubalansuota duomenų problema, kai įprastos klasės tinklo duomenų yra daug daugiau nei anomalijos klasės atvejų.

Klasteriais pagrįsti NADS metodai turi keletą privalumų. Pirma, jie duomenų grupes grupuoja neprižiūrimai, o tai rodo, kad jiems nereikia pateikti klasių etikečių stebėjimams, o tai yra labai sunkus procesas, kad būtų užtikrintas teisingas duomenų žymėjimas tiek normaliu, tiek užpuolimo atveju. Antra, jie yra veiksmingi norint sugrupuoti didelius duomenų rinkinius į panašias grupes, kad būtų galima aptikti tinklo anomalijas, kurios sumažina skaičiavimo sudėtingumą, ir jos veikia geriau nei klasifikavimo metodai. Tačiau yra ir trūkumų. Vienas iš klasterizavimu grindžiamų NADS trūkumų yra tas, kad jo klasterių sudarymas labai priklauso nuo jo efektyvumo klasifikuojant įprastus atvejus, o kitas yra tas, kad dinamiškas teisėtų tinklo duomenų profilio atnaujinimas užima daug laiko. Galiausiai, priklausomybė nuo vienos iš trijų aukščiau esančių prielaidų kartais yra sudėtinga, norint veiksmingai atpažinti neįprastą elgesį, nes tai sukelia aukštą klaidingo aliarmo dažnį ir ypač išpuolių atvejai gali paslėpti save įprastoje grupėje.

1.5.3 Gilaus mokymosi metodai tinklo anomalijoms aptikti

Pagrindinė gilaus mokymosi metodų teorija yra pažangios dirbtinio neuroninio tinklo architektūros (žr. 1.3 pav.) (toliau įvardijama kaip ANN), kurią įkvepia žmogaus smegenys, panaudojimas visiškai kitaip nei tradiciniai skaitmeniniai metodai. ANN yra mašininio mokymosi algoritmai, kurie įvestis į išvestis paverčia dirbtiniu neuronų rinkinio apdorojimu. Šie metodai yra klasifikuojami į nuoseklų ir gilų mokymąsi. Nuosekliame neuroniniame tinkle dažnai yra vienas ar du paslėpti sluoksniai, tuo tarpu gilų tinklą sudaro keli paslėpti sluoksniai su keliomis architektūromis.



1.3 pav. Dirbtinio neuroninio tinklo schema

Pastaruoju metu gilaus mokymosi tinklai yra plačiai naudojami įvairių modelių atpažinimui ir tinklo aplikacijoms, nes jie gali išsamiai išmokti apskaičiavimo procesą. Taikant NADS metodiką, nuosekliems ir giliems tinklams reikia tam tikros informacijos apie leistiną duomenų klasę, kad būtų galima sistemingai keisti sujungimo neuronus, išmokti tinklo svorius ir gauti modelį, kuris galėtų atskirti išpuolius nuo įprasto elgesio. Giluminio mokymosi tinklai skirstomi į įvairius tipus, atsižvelgiant į jo architektūrinį dizainą, kurį sudaro netiesinio apdorojimo lygių hierarchiniai sluoksniai. Šie tinklai skirstomi į generacinę (angl. *generative*) ir diskriminacinę architektūras. Generacinė architektūra apskaičiuoja stebimų duomenų ir jų klasių tikimybių pasiskirstymus, kuriuos apima toliau išvardijami modeliai.

- Rekurentinis neuroninis tinklas (RNN) - tai prižiūrimas ir (arba) neprižiūrimas mokymosi modelis. Pagrindinė RNN teorija yra ta, kad informacija yra sujungta ilgomis sekomis per sluoksnių ryšį su grįžtamojo ryšio kilpa. Tarp jo sluoksnių yra nukreiptas ciklas, kuris padidina jo patikimumą, sukuriant vidinę atmintį ankstesnio įėjimo duomenims registruoti;
- „Deep Auto Encoder“ (DAE) - naudojamas efektyviam kodavimui išmokti neprižiūrimas. Paprasčiausia DAE architektūra apima įvesties sluoksnį, daugiau nei vieną paslėptą sluoksnį ir išvestinį sluoksnį, kurio įvesties sluoksnyje yra toks pat neuronų skaičius rekonstrukcijai;
- „Deep Boltzmann Machine“ (DBM) - tai netiesioginis tikimybių modelis, apimantis viso tinklo energiją ir stochastinius vienetus, siekiant gauti dvejetainius rezultatus. Paslėptų sluoksnių sumažinimui taikoma riboto naudojimo Boltzmann mašina (RBM), kuri neleidžia jungtis tarp sluoksnių tarp paslėptų elementų. Treniruodami krūvą DBM, naudodami nepaženklintus

duomenis kaip kito sluoksnio įvestį ir įdėdami sluoksnį diskriminacijai, gali būti sukurta DBN architektūra;
- „Dee Belief Network“ (DBN) - apima daugybę paslėptų sluoksnių, kai ryšys yra tarp sluoksnių, o ne tarp kiekvieno sluoksnio vienetų. Tai neprižiūrimų ir prižiūrimų mokymosi tinklų rinkinys. Neprižiūrimas modelis išmokstamas naudojant nemandagų ryšį kiekviename lygmenyje, tuo tarpu prižiūrimas tinklas yra vienas ar keli sluoksniai, sujungti užduotims klasifikuoti.

Diskriminacinė architektūra įvertina vėlesnį klasių pasiskirstymą, atsižvelgiant į stebėtus duomenis, kuriuos sudaro RNN ir konvoliucinius neuroninis tinklas (CNN), aptariamas toliau.

- RNN - naudojasi diskriminacine galia klasifikavimo užduotims atlikti, ir tai įvyksta, kai modelio išvestis žymima seka su įvestimi;
- konvoliucinis neuroninis tinklas (CNN) - tai erdvėje nekintamas multiperceptroninis ANN, biologiškai įkvėptas organizuojant gyvūnų regos žievę. Jis turi daug paslėptų sluoksnių, kuriuos paprastai sudaro konvoliuciniai sluoksniai, jungiamieji sluoksniai, visiškai sujungti sluoksniai ir normalizavimo sluoksnis. Konvoliuciniai sluoksniai turi daugybę svorių, kurie turi mažus parametrus, todėl CNN mokymo procese yra lengvesnis, palyginti su kitais modeliais, turinčiais tą patį paslėptų sluoksnių skaičių.

Daugybė mokslinių tyrimų neseniai pritaikė gilaus mokymosi metodus NADS. Buvo panaudota DBN pagrindu sukurta NADS, sukonfigūruojant gobšą mokymosi algoritmą, sudarytą iš sluoksnių po sluoksnio, kad būtų galima išmokti kiekvieną RBM krūvą vienu metu, norint aptikti įsibrovimo įvykius. Vėliau buvo sukurta gilaus automatinio kodavimo technika, siekiant sumažinti duomenų matmenis, kurie buvo laikomi pradiniu etapu klasifikuojant tinklo stebėjimus. Negilus ANN algoritmas buvo naudojamas kaip klasifikatorius, norint įvertinti DAE efektyvumą. Dar vėliau pasiūlyta RNN pagrįsti NADS sistemos, kurios atpažintų kenksmingus tinklo egzempliorius.

Pastebėta, kad giluminiai mokymosi algoritmai galėtų žymiai pagerinti NADS veikimą, turėdami aukštą aptikimo tikslumą ir žemą klaidingų aliarmų dažnį. Tačiau jie paprastai užima daug laiko tinklo duomenims apdoroti, kad nustatytų geriausius neuroninius svorius, kad būtų kuo labiau sumažintos klasifikavimo klaidos.

1.5.4 Žiniomis paremti metodai tinklo anomalijoms aptikti

Žiniomis pagrįsti metodai sukuria modelių rinkinį iš pradinių duomenų, kad būtų galima klasifikuoti duomenų taškus atsižvelgiant į klasės etiketes. Tinklų srauto duomenys tiriami atsižvelgiant į iš anksto nustatytus išpuolių modelius ir sistemos pažeidžiamumas aptikti kenksmingus įvykius bei kelti aliarmą. Nors šie metodai gali atpažinti žinomus išpuolius, jie negali nustatyti nulinės dienos išpuolių, išskyrus atvejus, kai profilis sudaromas pagal įvairius normalius modelius, kaip NADS.

Įprasti žiniomis pagrįsti NADS metodai yra pagrįsti taisyklėmis ir ekspertais, taip pat ontologija ir logika. Taisyklėmis pagrįsti metodai modeliuoja surinktas žinias apie įtartinus tinklo įvykius, kurie leidžia naršyti tinklo srauto duomenis, kad būtų galima rasti esamų pažeidžiamumų įrodymų. Ekspertų sistemą sudaro taisyklės, apibrėžiančios atakos įvykius, kai tinklo srauto duomenys paverčiami modeliais pagal jų santykinį svorį sistemoje, o išvadų variklis suderina iš anksto nustatytas taisykles su dabartine sistemos būkle, kad aptiktų išpuolių veiksmus. Taisyklėmis paremti ir ekspertų sistemos metodai buvo plačiai taikomi aptikti įtartinus tinklo įvykius, o ontologija ir logika pagrįsti įsibrovimo parašų modeliai, pagrįsti logine struktūra, įtraukiant tinklo srauto duomenų apribojimus ir statistines charakteristikas.

- „Snort“ („Snort“ įrankis) yra viena iš populiariausių taisyklėmis pagrįstų ir atviro kodo įsilaužimų aptikimo sistema. Jos taisyklės atpažįsta kenksmingus tinklo paketus, suderindamos dabartinį paketą su iš anksto nustatytomis taisyklėmis ir negali aptikti nulinės dienos atakų, tačiau sukuria aukštą FPR dėl savo metodikos, skirtos identifikuoti išpuolių parašus. Šiuo metu „Snort“ apima daugiau nei 20 000 taisyklių, kurias paprastai atnaujina vartotojai;
- „Petri“ tinklų įrankis, kuris buvo sukurtas Purdue universitete, yra žiniomis pagrįstas IDS, kurį sudaro nukreipti dvipusiai grafikai ir spalvoti Petri tinklai (CPN), vaizduojantys įsilaužimo parašus, pavyzdys. Šis įrankis buvo naudojamas kuriant įsibrovimų aptikimo mūsų laikais (IDIOT) įrankį netinkamam naudojimui aptikti. Nors šis įrankis gali lengvai atvaizduoti mažus tinklo duomenis ir padeda atpažinti žinomas atakas, jo procesą, skirtą atakų parašui suderinti su iš anksto nustatytomis taisyklėmis, labai sunku vykdyti realioje tinklo aplinkoje ir tai užtrunka ilgą apdorojimo laiką.

Praeityje buvo pasiūlytas įsibrovimų nustatymo įrankis, kuris atpažįsta kenksmingą statistinį elgesį, nustatydamas taisyklių rinkinį, statistiškai vaizduojantį vartotojų elgesį, naudojantis tam tikru laikotarpiu jų veiklos žurnalus. Tada jis suderina esamą veiklą su saugomomis taisyklėmis, kad būtų galima nustatyti įtartiną elgesį.

Žiniomis grindžiami NADS mechanizmai turi tam tikrų pranašumų: pirma, jie yra pakankamai tvirti ir lankstūs, kad diskriminuotų esamas mažų tinklo srauto duomenų atakas; ir, antra, pasiekti aukštą aptikimo procentą, jei galima tinkamai išgauti didelę žinių apie teisėtus ir anomalius atvejus bazę. Tačiau šie metodai negali nustatyti retų ar nulinės dienos anomalijų. Galiausiai, jų dinaminio taisyklių atnaujinimo procedūros yra labai sudėtingos, o jų apdorojimo laikas yra labai brangus kuriant NADS sistemą.

1.5.5 Deriniais paremti metodai tinklo anomalijoms aptikti

Derinių pagrindu sukurta metodika naudoja įvairius mechanizmus duomenų taškų efektyviam ir efektyviam klasifikavimui. Grupių mokymosi metodai integruoja daugelį metodų ir juos sujungia, kad būtų pasiektas bendras tikslumas, kuris pranoksta kiekvieno klasifikatoriaus tikslumą. Pirma, grupė pagerina aptikimo tikslumą sukuriant patobulintą sudėtinį klasifikatorių, kuris sujungia anksčiau naudotų klasifikavimo metodų išvadas į vieną prognozatorių. Antra, inkrementinės grupės kūrimas, mokantis klaidingai klasifikuoti stebėjimus, įgytus pagal ankstesnį modelį. Trečia, grupės apibendrinimas išgauna didžiausią apibendrintą tikslumą, panaudodamas kiekvienos klasės tikimybes iš tam tikro klasifikavimo algoritmo.

Sintezės metodais pagrįsti metodai, integruojantys skirtingų klasifikatorių priimamus sprendimus, išryškėjo kaip metodai, galintys sustiprinti galutinį sprendimą, jų taksonomija susidedanti iš trijų lygių: duomenų, ypatybės ir sprendimo. Kai kurie metodai išsprendžia didelio duomenų kiekio problemą priimdami tik svarbius požymius, o kiti sujungia įvairius požymius klasifikuojančius metodus, naudodami hierarchinius abstrakcijos lygius arba naudojamų požymių tipus.

Deriniais pagrįsti metodai yra naudingi, nes jais pasiekiamas didesnis tikslumas ir aptikimo dažnis nei su pavieniais, tuo pačiu reikalaujant kai kurių parametrų, kuriuos galima tiksliai pakoreguoti. Tačiau sunku priimti nuoseklius ir nešališkus klasifikavimo metodus, nes juos sujungti reikia naudojant hibridizacijos priemonę. Taip pat akivaizdu, kad jų skaičiavimo išlaidos didžiuliams tinklo paketų kiekiams yra aukštos dėl naudojamų klasifikatorių skaičiaus.

1.5.6 Statistiniai metodai tinklo anomalijoms aptikti

Žiūrint iš statistinio aspekto, anomalija yra retas įvykis, atsirandantis tarp natūralių duomenų įvykių ir matuojamas statistiniais metodais, kurie gali būti pirmosios eilės, tokie kaip vidurkiai ir standartiniai nuokrypiai, antrosios eilės, tokie kaip koreliacijos matavimai, arba trečios eilės, tokie kaip hipotezės tyrimas, mišinių modeliai ir išvadų metodai. Tinklo anomalijų aptikimo sistemose šie metodai tinka statistiniam teisėtų tinklo duomenų modeliui ir tada naudojami statistiniai testai, naudojant ribinę / bazinę ar tikimybės sąlygą, siekiant nustatyti nukrypusius atvejus kaip anomalijas. Statistiniu pagrindu pagrįsti metodai yra klasifikuojami kaip neparametriniai ir parametriniai, kurie abu buvo plačiai taikomi kuriant statistinius NADS modelius.

1.5.6.1 Neparametriniai metodai

Neparametriniai metodai nedaro jokių prielaidų dėl pateiktų duomenų statistinių charakteristikų. Jie sukuria modelį, kai jie veikia, ir bando išspręsti duomenų sudėtingumą, kad efektyviai pritaikytų duomenų taškus. Vienas iš paprasčiausių neparametrinių statistinių metodų yra histogramų įrankių naudojimas, grafiškai parodantis lentelių duomenų dažnį. NADS sukuriama normali histograma ir tada nustatomi nauji išbandyti duomenų taškai, kurie, jei jie nepatenka į įprastą histogramą, laikomi anomaliais atvejais. Daug kintamųjų turintiems tinklams naudojamas elementų lygio histogramos, o bendras bandymo duomenų taškas gaunamas kaupiant pasirinktų funkcijų balus. Dažniausiai naudojami neparametriniai metodai yra šie:

- **branduolio tankio įvertis** – neparametrinis metodas, kuris grindžia savo skaičiavimus kai kuriais branduolio pasiskirstymais, tokiais kaip Gauso, visiems mėginio vietos duomenims, tada integruoja visų paskirstymų vietinius indėlius. Kiekvieno mėginio tikimybės tankio įvertinimas priklauso nuo duomenų taškų, esančių lokalizuotoje branduolio kaimynystėje;
- **neigiama atranka** buvo plačiai taikoma aptikti anomalius tinklo atvejus. Neigiamos atrankos teoriją įkvėpė žmogaus imuninės sistemos savybės, pagal kurias galima identifikuoti antigenus, tai reiškia, kad galima aptikti bet ką, kas nėra žmogaus kūno dalis, pavyzdžiui, virusus ir bakterijas. Atakos aptikimo pagrindinis tikslas yra atskirti „save“, kurie primena įprastą stebimos sistemos veikimą, ir nesavanaudiškus duomenis, rodančius nenormalius duomenis.

1.5.6.2 Parametriniai metodai

Parametriniai metodai daro prielaidą, kad tinklo duomenys seka tam tikru pasiskirstymu, pavyzdžiui, kad Gauso paskirstymas įvertina pateiktų duomenų parametrus. Tačiau realiame tinkle nėra žinomas pagrindinis tinklo srauto duomenų paskirstymas, svarbu nurodyti, koks tikimybės pasiskirstymas gali atitikti duomenis su palyginti mažu klaidų lygiu.

Pastebėta, kad tinklo duomenys nepriklauso Gauso paskirstymui, naudojant Kolmogorovo-Smirnovo (KS) bandymo metodą geriau taikyti ne Gauso pasiskirstymą, tokius kaip Gauso Mišinio modelis, Beta mišinio modelis (BMM) arba Dirichlet mišinio modelis (DMM), tinklo duomenims. Šių paskirstymų tikimybių tankio funkcijos turi būti modeliuojamos pagal įeinančius tinklo duomenis, iš kurių dinamiškai turėtų būti pritaikyti jų parametrai, užuot nustatę statinį parametrą. Taip sukuriamas lankstus modelis, išskiriantis anomalijas nuo įprastų stebėjimų. Toliau aptariama dažniausiai naudojamų parametrinių metodų metodika.

1. **Dalelių filtras** – išvesties mechanizmas, matuojantis nežinomą būseną nuo tam tikro stebėjimo rinkinio tam tikru laiku, o galinį pasiskirstymą nustato svertinių dalelių rinkinys.

2. **Bajeso tinklai (BN)** – grafinis tikimybių pasiskirstymas, priimant sprendimus dėl neaiškių duomenų. Pavyzdžiui, naudojant PCA, buvo sukurtas BN NADS, kuris apskaičiavo PCA aukščiausius reitingus pasižyminčias savybes ir pasirinktas savybes bei jų komponentus naudojo kaip svorį, siekdami patobulinti tradicinę Bajeso techniką. Eksperimento rezultatai parodė, kad tai galėtų efektyviai sumažinti duomenų matmenis ir pagerinti aptikimo tikslumą;
3. **Baigtinio mišinio modelį** galima apibrėžti kaip išgaubtą dviejų ar daugiau tikimybių pasiskirstymo funkcijų rinkinių derinį, kurio jungtinės savybės gali apytiksliai suderinti bet kokį savavališką paskirstymą, tai yra galingas ir lankstus tikimybinis modelio įrankis, skirtas vieno faktoriams ir keleto faktorių duomenims. Tinklo duomenys paprastai laikomi daugiapakopiais, nes jie turi d dimensijas, kad būtų galima atskirti išpuolius ir įprastus atvejus. Mišinių modeliams reikia daug įprastų egzempliorių, kad būtų galima teisingai įvertinti jų parametrus, ir sunku pasirinkti tinkamą slenkstį, kaip ir 3 lygtyje, kuri atakų atvejus skiria nuo įprastos treniruočių klasės su tam tikru balu;
4. **Logistinė regresija** – klasifikavimo algoritmas, naudojamas, kai tikslinio kintamojo vertė yra kategoriško pobūdžio. Logistinė regresija dažniausiai naudojama, kai nagrinėjami duomenys turi dvejetainę išvestį, tai reiškia kad yra tik 2 galimos išvestys, kai ji priklauso vienai ar kitai klasei arba yra 0 arba 1. Svarbu suprasti, kad logistinė regresija turėtų būti naudojama tik tada, kai tiksliniai kintamieji patenka į atskiras kategorijas. Logistinė regresija neturėtų būti naudojama, jei yra nenutrūkstamų reikšmių diapazonas.

1.5.7 Medžių struktūros metodai tinklo anomalijoms aptikti

1.5.7.1 Sprendimų medis

Sprendimų medžio struktūra yra paprasta – ji hierarchiškai suskirsto duomenis į pogrupius, kurie vėliau vėl padalijami į mažesnes skaidinius ar šakas, kol tampa „gryni“, o tai reiškia, kad šakos viduje esančios savybės priklauso tai pačiai klasei. Tokios klasės vadinamos „lapais“. Daugiau apie sprendimų medžius galite perskaityti čia. Iš esmės medžių klasifikacija turi tokį srautą:

- viršutinis mazgas reiškia šaknį;
- vidinis mazgas atspindi ypatybes;
- lapai atspindi rezultatą.

1.5.7.2 Atsitiktinių miškų metodas

Atsitiktinis miškas, kaip rodo jo pavadinimas, susideda iš daugybės individualių sprendimų medžių, kurie veikia kaip vienas. Kiekvienas atskiras medis atsitiktiniame miške pateikia klasės prognozę, o klasė, surinkusi daugiausiai balsų, tampa mūsų modelio prognoze.

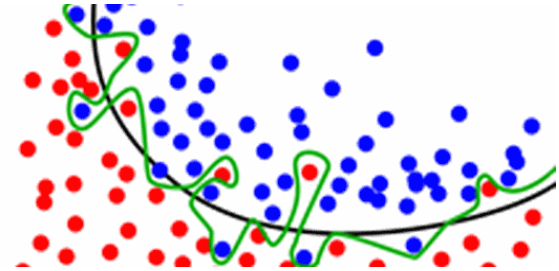
Pagrindinė atsitiktinio miško samprata yra paprasta – minios išmintis. Kalbant apie duomenų mokslą, atsitiktinio miško modelis veikia gerai, nes didelis kiekis modelių (medžių), veikiančių kaip vienas, pranoks bet kurį atskirą modelį.

1.6 Tinklo anomalijos jautrumo problema

Kadangi tinklo anomalijoms naudojami save apsimokantys metodai, todėl juose naudojami apsimokymo metodai sudaro didelę dalį anomalijos aptikimo sistemoje. Jei yra naudojami duomenys kuriuose nėra jokių pašalinių tinklo srauto parametrų, tokia sistema pasidaro per daug jautri visiems anomalijos požymius turintiems įvykiams. Tai galime pamatyti 1.4 pav.. Darbe [11] rašoma, kad modelis išmoko tvirtai parodyti normalią klasę žemo matmens latentinėje erdvėje ir rekonstruoti ją į

pradinę erdvę, tuo pačiu tapdamas nejautrus mokymo anomalijoms. Abiejų klasių rekonstravimo klaidos iš dalies sutampa, tačiau padidėjęs atskyrimas lemia didesnį subalansuotą tikslumą. Todėl naudojant tam tikrus metodų apsimokymui skirtus tinklo srauto duomenis turime žinoti ar jis yra su „triukšmu“ ar ne.

Ši problema aprašoma ir [12] šaltinyje. Jame pažymima, kad apmokymo duomenyse esančias anomalijas, jau apmokytas, metodas aptinka daug sunkiau. Be to galima ir persimokymo problema, kuomet metodas neaptiks anomalijos, kuri skirsis ir labai mažais požymiais nuo mokymosi procese turėtos anomalijos.



1.4 pav. Persimokymo situacija neuroniniuose tinkluose

1.7 Tinklo anomalijų aptikimo metodų tyrimas ir apžvalga

SVM klasifikatorius buvo tiriamas šaltinyje [7]. Šiame tyrime buvo naudotas „KDD CUP 99“ duomenų rinkinys, kuriame klasifikuojami keturių tipų atakų įrašai (DoS, zondavimo, R2L ir U2R). Šis rinkinys tirtas NIDS sistemos analizuojamų parametrų principu. Duomenų rinkinio patvirtinimo tikslumas yra 89,85%, o sumažinus tiriamų duomenų dydį iki 10%, buvo gautas 99,9% tikslumas.

KNN metodas buvo tirtas darbe [8]. Šiame darbe šis metodas buvo tiriamas naudojant NSL-KDD duomenų rinkinį. Šiame darbe buvo ištobulintas KNN metodas, kuris su darbe tirtomis 10 svarbiausių rinkinio savybių. Metodas gauna 88,21% tikslumą.

DBN klasifikatorius naudojant „KDD CUP 99“ duomenų rinkinį, darbe [9], tiriant skirtingas tinklo struktūras gauna rezultatus svyruojančius nuo 75,6% iki 92,33%.

Sprendimų medis buvo tiriamas darbe [10]. Šiame darbe metodas buvo tiriamas naudojant NSL-KDD duomenų rinkinį. Rezultatas tikrinamas pagal penkių klasių klasifikaciją, sumažinant savybes ir kiekvieno genėjimo lygio rezultatą. Buvo gauti rezultatai tiriant 5 klases ir 2 klases. 5 tirtų klasių rezultatai rodo 83,66% tikslum1, o tiriant 2 klases – 90,31%.

Šaltinyje [13] aprašyta dabar naudojamų metodikų analizė (matoma 1.1 lentelė.).

1.1 lentelė. Esamų anomalijos aptikimo metodikų analizė ir palyginimas.

Autorius	**Metodo tipas**	**Algoritmas**	**Pliusai ir minusai**
Gharbaoui et al.,2013	Statistinis	Nuoseklus hipotezės tyrimas	Aptikimo galimybės leidžia pasiekti gerą rezultatą, sumažinti klaidingą aliarmą, išbandyti realiu laiku
Nguyen & Roughan, 2013	Statistinis	Paslėptas Markovo modelis	Mažos skaičiavimo ir ryšių pridėtinės išlaidos, tinka IPT, tik mažo dydžio duomenims

Fernandes et al., 2016	Statistinis	PCA + skruzdžių kolonijos	Geba nustatyti anomalų elgesį, Skaičiavimas labai sudėtingas
Parwez et al.,2017	Klasterizavimas	k-reikšmių ir hierarchinis grupavimas + neuroninis tinklas	Mažas k-klasterių sudėtingumas, geresnis našumas atliekant hierarchinį grupavimą, hierarchinis grupavimas susiduria su erdvės sudėtingumu dideliam duomenų rinkiniui
Zhu, C. et al., 2015	Klasifikatorius	Bajeso tinklas	Labai efektyvus norint nustatyti anomaliją, reikalauja vartotojo įsikišimo (eksperto), kad pritaikytų pasikeitusias tikimybes
Z. Zhang et al., 2016	Gilaus mokymosi	Informacijos entropija + neuroninio tinklo atgalinis plitimas	Gali pagerinti sistemos stabilumą, dinamiškai prisitaikyti prie eismo pokyčių, sumažinti klaidingą aliarmo dažnį, nustatyta, kad entropijos vertė yra per jautri
Shabtai et al., 2010	Žiniomis pagrįstas	žiniomis grįstas laikinasis	Palaikant piktnaudžiavimo ir anomalijos nustatymą, KBTA buvo pritaikytas mobiliesiems įrenginiams, kurių ištekliai yra riboti
Usman et al., 2015	Klasifikatorius	Miglota logika (angl. *Fuzzy logic*)	Didelis tikslumas nustatant tarpsluoksnio anomalijas, mažas energijos suvartojimas, reikalingos pradinės srities žinios, nepatikimas perduoti mobilųjį agentą (esant blogam ryšiui)
Alipour, H. Et al.,2015	Save apsimokantis (prižiūrimas)	n-grams	Sistema gali aptikti skirtumų ataką, aukštą aptikimo dažnį, žemas klaidingų aliarmų kiekis, dirba su iš anksto paženklintais duomenimis
Dromard et al., 2017	Save apsimokantis (neprižiūrimas)	tinklelio grupavimo algoritmas	Didelis aptikimo greičio efektyvumas, mažas klaidingas aliarmas, greitis turi būti pagerintas
Fawzy et al., 2013	Artimiausio kaimyno	K-artimiausio kaimyno	Geba aptikti pašalines reikšmes, gali klasifikuoti triukšmingus duomenis ar įdomius įvykius, nepatikrintas didesniame duomenų rinkinyje
Oreilly et al., 2016	Dekompozicinis	PCA	Gebėjimas sumažinti matmenis, skaičiavimas yra sudėtingas
Erfani et al.,2016	Kombinacinis	Gilaus pasitikėjimo tinklas + vienos klasės SVM	Veiksmingas, tikslus ir adaptyvus anomalijos aptikimas. Gali būti įdiegtas naudojant didelio masto ir didelius matmenis. išbandytas tik jutiklių tinklo duomenų rinkiniuose, todėl negarantuojama kitam domenui.

Tinklai su įvairiomis programomis ir įranga generuoja didžiulį kiekį duomenų tiek skaičiaus, tiek tipo. Tai susiję su duomenų matmenimis. Kaip žinome, dimensijų dydis yra viena iš problemų nustatant anomalijas.

Tiek daug tyrėjų naudojamų anomalijų aptikimo tyrimuose rezultatas yra teigiami ir neigiami šių metodų požymiai. Remiantis geriausiomis šios apklausos autorių žiniomis, populiariausios anomalijų aptikimo problemos apima ir neapsiriboja aptikimo galimybėmis, tokiomis kaip aptikimo dažnis ir klaidingo aliarmo dažnis. Kita problema, susijusi su aptikimo galimybėmis, yra matmenų sumažinimas ir skaičiavimo sudėtingumas. Kai kurie tyrėjai susirūpinę dėl skaičiavimo laiko ir plečiamumo. Iš šio tyrimo padaryta išvada, kad kai kuriais metodais buvo pasiekti puikūs aptikimo gebėjimai, tačiau yra tokių pasekmių kaip didelis klaidingas pavojaus signalas, skaičiavimo sudėtingumas, skaičiavimo laikas.

1.8 Esamos tinklo anomalijų aptikimo sistemos

1.8.1 Komercinės anomalijų aptikimo sistemos

Buvo paimtos 4, jau esamos, tinklo anomalijų aptikimo sistemos: „Cisco Stealthwatch“ [14], „IBM QRadar“ [15], „McAfee Network Threat Behavior Analysis“ [16], „Lastline DefenderTM“ [17].

„Cisco® Stealthwatch“ yra pramoninio lygio saugumo analizės sprendimas, užtikrinantis suprantamą grėsmių matomumą išplėstame tinkle. Ši sistema gali aptikti pažangias grėsmes ir į jas reaguoti bei padėti supaprastinti tinklo segmentus, naudojant elgesio modeliavimo, daugiasluoksnio mašininio mokymosi ir visuotinės grėsmės informacijos derinį. Kadangi užpuolikai nenaudoja tik vieno būdo, kad pažeistų jūsų tinklą, „Stealthwatch“ naudoja kelis analizės metodus, kad anksti nustatytų grėsmes, ir padeda užtikrinti, kad jos būtų izoliuotos.

Atakos taktikos raida kartu su prastu grėsmės matomumu palaiko gynėjus ant kojų, ypač kai priešininkai išnaudoja vartotojus ir naudoja individualiai sukurtas, trumpai pritaikytas kenkėjiškas programas, kad nustatytų savo pradinę poziciją. Todėl saugumo analitika, kuri kaupia saugumo duomenis ir paverčia juos veiksminga grėsmės įžvalga, tampa saugumo komandų prioritetu. Tūkstančiai apsaugos komandų visame pasaulyje diegė sprendimą, kad automatiškai nustatytų, apimtų grėsmes ir nustatytų jų prioritetus. Darbe [15] paaiškinta, kaip „IBM QRadar“ renka ir analizuoja duomenis, kad saugos komandos galėtų geriau nustatyti ir valdyti grėsmes.

McAfee®“ tinklo grėsmės elgsenos analizės (NTBA) virtualus prietaisas yra integruotas „McAfee®“ tinklo saugumo platformos komponentas, užtikrinantis tinklo infrastruktūros matomumą realiu laiku ir apsaugą nuo grėsmių. Analizuodamas srautą iš jungiklių ir maršrutizatorių, „McAfee NTBA“ nustato rizikingą elgesį tinkle ir veiksmingai apsaugo nuo slaptų atakų. Sistema visapusiškai įvertina tinklo lygio grėsmes ir nustato bendrą kiekvieno tinklo elemento elgesį, įskaitant kenkėjiškas programas, nulinės dienos atakas, „Botnet“ tinklus ir kirminus. NTBA taip pat teikia „McAfee“ tinklo saugumo platformos pažangių variklių komponentus, įskaitant realaus laiko emuliacijos variklį, kuris identifikuoja kenkėjiškas programas be parašų.

Tinklo aptikimo ir reagavimo (NDR) platforma „Lastline DefenderTM“ aptinka sudėtines grėsmes ir jose pateikia prieš tai, kai jos sutrikdo jūsų įmonės tinklą. Tai pateikia aukščiausią kibernetinio saugumo pramonės įžvalgą apie pažangias grėsmes, kylančias ar veikiančias jūsų vietiniame ir debesų tinkle, leidžiančią jūsų saugos komandai greičiau ir veiksmingiau reaguoti į grėsmes. „Defender“ platformoje naudojamos trys papildomos dirbtinio intelekto palaikomos technologijos, skirtos aptikti

pažangias grėsmes, kurių praleidžia kitos priemonės, ir žymiai sumažinti klaidingus teiginius: elgesio analizė siekiant aptikti kenksmingą turinį, bandantį patekti į jūsų tinklą žiniatinkliu ar el. paštu, tinklo srauto analizė (NTA), siekiant aptikti šoninį vengtinų grėsmių judėjimą jau jūsų tinkle, įsibrovimo aptikimas / prevencija (IDPS) žinomoms grėsmėms nustatyti. Šis unikalus derinys įgalina deterministinius aptikimus ir pašalina daugiausiai klaidingų teiginių. Galite reaguoti greičiau ir efektyviau, turėdami mažiau išteklių.

1.8.2 Atviro kodo tinklo anomalijų aptikimo sistemos

Taip pat tinklo anomalijų aptikimui yra sukurtų ir atviro kodo programų, kurias visi vartotojai gali naudoti ir tobulinti. Šiuo metu populiariausios atviro kodo priemonės anomalijų aptikimui yra pateiktos žemiau.

„Ourmon“ [18] yra statistiškai orientuota atvirojo kodo tinklo stebėjimo ir anomalijų aptikimo sistema. Tai taip pat gali būti vertinama kaip srautų surinkimo sistema. Zondas surenka svarbiais laikomus paketus ir siunčia vidinius apibrėžimus atgal į grafikos ekrano sistemą, kuri gali būti tame pačiame pagrindiniame kompiuteryje. „Ourmon“ naudoja labiau ekstremalų apibendrinimą nei tinklo srautą. „Ourmon“ nerenka visų paketų, nes vienas pagrindinių projektavimo tikslų yra išgauti signalą iš triukšmo, o ne laikyti visą triukšmą milžiniškame krepšyje, darant prielaidą, kad jį galėsite suvokti „vėliau“. „Ourmon“ taip pat turi savo sąvokų „tuples“ sąvoką ir, nors ji palaiko tradicinį srauto rinkinį,

Atvirojo kodo sistema „Tripwire“ [19] yra pagrindinio kompiuterio įsibrovimo aptikimo sistema, skirta aptikti failų sistemos objektų pokyčius. Pirmą kartą inicijavęs „Tripwire“, ji nuskaito failų sistemą, kaip nurodo sistemos administratorius, ir saugo kiekvieno failo informaciją duomenų bazėje. Pakeitus failus ir atliekant būsimus nuskaitymus, rezultatai lyginami su išsaugotomis vertėmis, o apie pasikeitimus pranešama vartotojams. „Tripwire“ naudoja kriptografines maišas, kad aptiktų failų pokyčius. Be nuskaitymo failų pakeitimų, jis naudojamas ir vientisumui užtikrinti, pokyčių valdymui ir politikos laikymuisi užtikrinti.

„Security Onion“ [20] yra „Linux“ versija, skirta įsilaužimui aptikti, tinklo saugumo stebėjimui ir žurnalų valdymui. Atvirojo kodo sistema pagrįsta „Ubuntu“ ir apima daugybę IDS įrankių, tokių kaip „Snort“, „Suricata“, „Bro“, „Sguil“, „Squert“, „Snorby“, ELSA, „Xplico“, „NetworkMiner“ ir daugelį kitų. „Security Onion“ suteikia didelį tinklo srauto, perspėjimo ir įtartinos veiklos matomumą ir kontekstą. Bet tam reikia tinkamo sistemos administratoriaus valdymo, kad peržiūrėtų perspėjimus, stebėtų tinklo veiklą ir reguliariai atnaujintų IDS pagrįstas aptikimo taisykles.

1.9 Analizės išvados

Atlikus mokslinių darbų tyrimą, buvo pastebėta jos anomalijų aptikimo tinkle problema vis dar yra aktuali. Kiekviena su internetiniais tinklais susijusi įmonė nori žinoti kas vyksta jų vidiniame ar tiekiamame tinkle, todėl jie pasitelkia tinklo anomalijų aptikimo sistemas.

Norint sužinoti kas yra normalus tinklo veikimas reikia giliai išanalizuoti savo tiriamo tinklo srautus, tam reikalingi tinklo srautų registratoriai, kurie surenka reikalingus duomenis. Populiariausias – „Cisco Netflow“.

Išnagrinėti 7 dažniausiai naudojamos anomalijų aptikimo ir klasifikavimo metodų grupės, kurios paremtos apsimokymu pagal sužymėtą srauto įrašų rinkinį. Tačiau trūksta daugumos metodų tikslumo palyginimo konkrečių atakų aptikimui ir kaip jį įtakoja apsimokymo imties dydis.

Iš analizuotos medžiagos matyti, kad metodų efektyvumas priklauso nuo apsimokymui pateikiamo įrašų rinkinio, jame esančio triukšmo ir imties dydžio, tačiau bendrų taisyklių nėra. Konkretūs įvertinimai gaunami tik eksperimentuojant.

Apžvelgtos komercinės ir atviro kodo anomalijų aptikimo ir atakų identifikavimo sistemos, tačiau nerasta jų efektyvumo įvertinimo konkrečių atakų aptikimui.

Pagal surinktą ir išanalizuotą medžiagą galima daryti išvadą, kad nėra vieningos aprašytų metodų vertinimo metodikos. Todėl tikslinga eksperimentiškai ištirti jų efektyvumą konkrečių atakų atveju ir tikslumo priklausomybę nuo apsimokymo imties dydžio.

2 Mašininio mokymo metodų taikymo atakoms aptikti naudojant *Cisco Netflow* tinklo įrašus projektas

2.1 Darbo tikslas ir keliami reikalavimai

Šio projekto tikslas suprojektuoti tinklo anomalijų aptikimo metodą, kuris naudotų „Cisco NetFlow“ tinklo registratoriaus gaunamus duomenis.

2.1.1 Reikalavimai duomenims

Tyrimui bus naudojami duomenų rinkiniai iš viešų duomenų bazių, įvertinant jų atitikimą „Cisco NetFlow“ tiesiogiai renkamais srauto požymiais ar išvestiniams įverčiams. Duomenų įrašo faile, privalo būti sužymėti išrašų tipai.

2.1.2 Funkciniai ir nefunkciniai reikalavimai

Funkciniai reikalavimai:

- galimybė tyrinėti atskirus metodus ir jų kombinacijas;
- galimybė sumažinti duomenų savybių kiekį atrenkant tik reikšmingas duoto tipo atakoms;
- aptikti DoS tipo atakas;
- aptikti U2R tipo atakas;
- aptikti R2L tipo atakas;
- aptikti zondavimo tipo atakas.

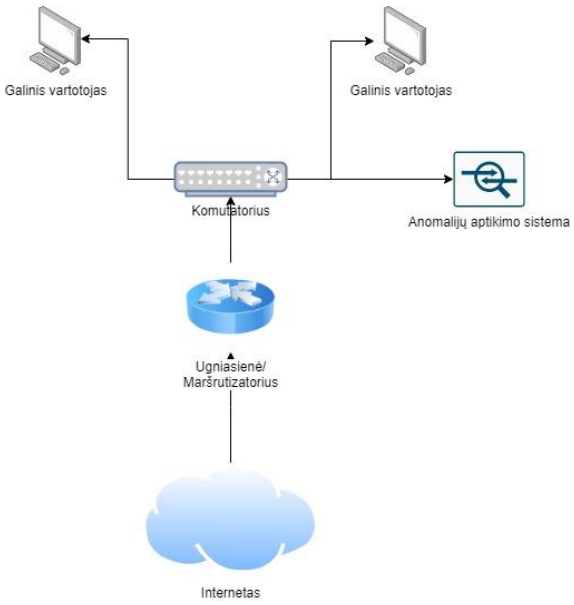
Nefunkciniai reikalavimai:

- Metodo apmokymas ir rezultatų išvedimas neturi viršyti 10 min.

2.1.3 Kokybės kriterijai

Metodo kokybė priklausys nuo anomalijų klasifikatorių tikslumo vertinimo tinklo sraute, kurios bus lyginamos su kitų tinklo anomalijų aptikimo metodų efektyvumo vertinimais.

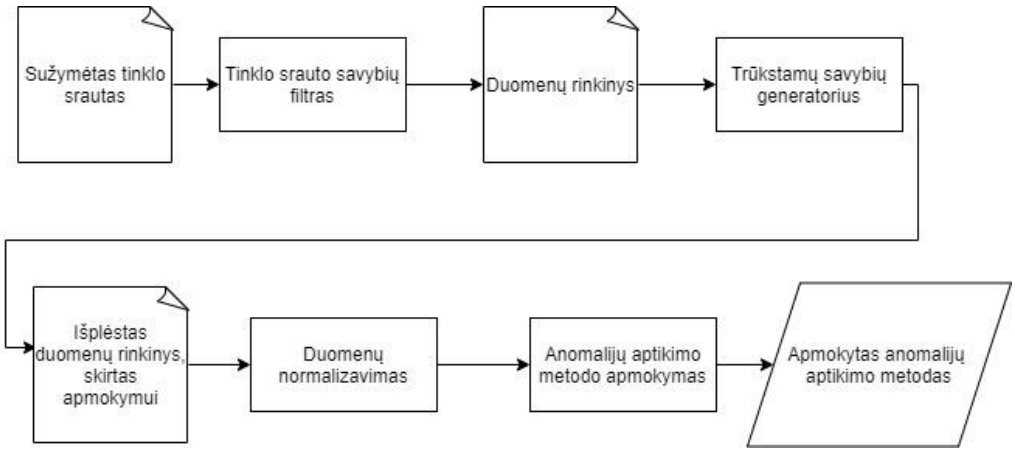
2.2 Anomalijų aptikimo modulio vieta tinklo architektūroje



2.1 pav. Anomalijų aptikimo modulio vieta tinklo architektūroje

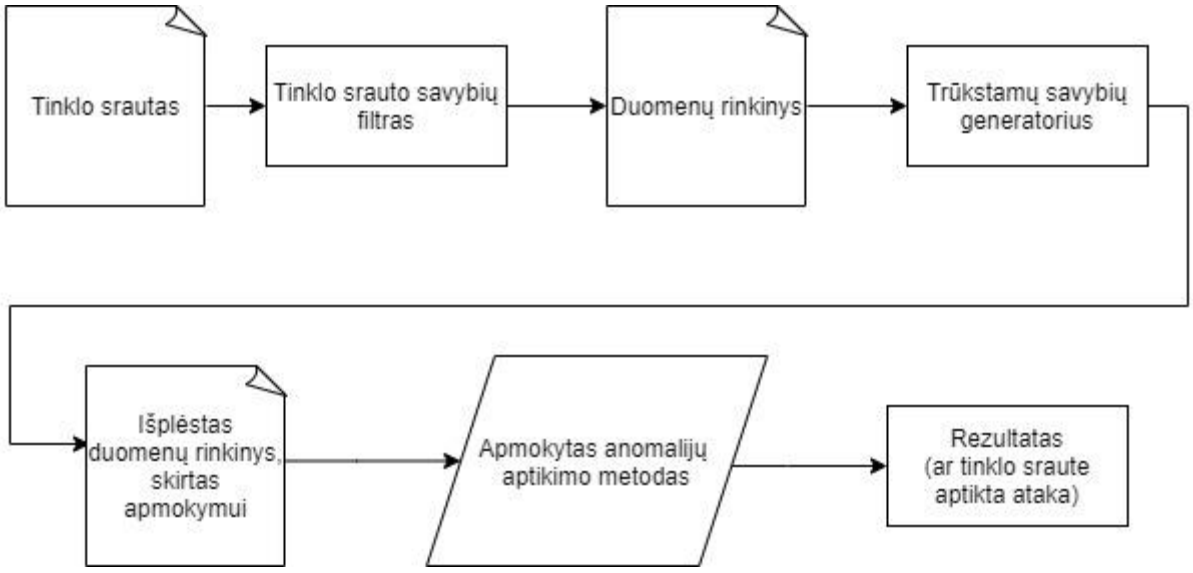
Teoriškai anomalijų aptikimo modulis turi būti analizuojamo tinklo viduje (žr. 2.1 pav.), t.y. už pagrindinio maršrutizatoriaus, tačiau anomalijų aptikimo modulis gali būti ir integruotas maršrutizatoriuje.

2.3 Sistemos architektūra



2.2 pav. Bendrinė tinklo anomalijų aptikimo sistemos apmokymo architektūra

Tinklo srauto anomalijų aptikimo sistema susideda iš pačio tinklo srauto įrašo, kuris pateikiamas filtrui. Jis atrenka tinklo srauto įrašus, kurie toliau analizuojami pagal tam tikras savybes ir išskiriamos svarbiausios. Tada atrenkama ir sugeneruojama trūkstamos savybės anomalijoms aptikti ir sukuriamas išplėstas duomenų rinkinys skirtas apsimokymui. Toliau atrinktas srautas patenka į anomalijų aptikimo metodą su kurio jis yra apmokomas. Šiuos procesus atvaizduoja 2.2 pav.



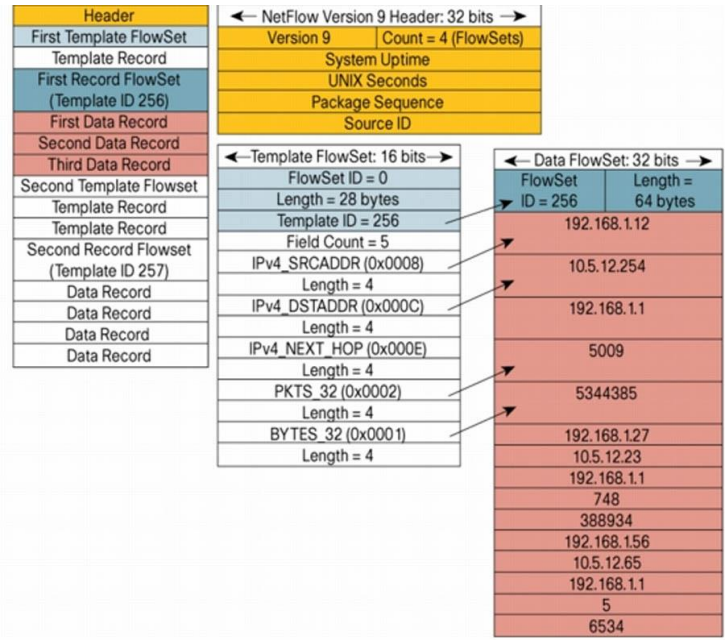
2.3 pav. Bendrinė tinklo anomalijų aptikimo sistemos architektūra

Anomalijų aptikimo metodas prognozuoja ar gyvame tinklo sraute šiuose tinklo duomenyse yra į anomalijas panašių srauto paketų ir pateikia rezultatus, pagal kuriuos tinklo administratorius nusprendžia ar imtis kažkokių veiksmų. Tai vykdoma pro tinklo srauto filtrą tokiu pačiu būdu praleidžiant norimą patikrinti tinklo srautą. Tai atlikus bus suvienodinami srautų duomenys, pagal kuriuos yra aptinkama ataka. Tai galime pamatyti 2.3 pav.

2.4 „Cisco NetFlow“ paketo struktūra ir atakų aptikimas jame

2.4.1 Apie „Cisco Netflow“

IT organizacijoms reikia intelektualių sprendimų, kurie yra plačiai paplitę, pagrįsti elgsena ir papildantys dabartinius zonos saugos sprendimus. Vienas iš tokių sprendimų yra naudoti pačią tinklo infrastruktūrą, kad veiktų kaip jutiklis. Tai daroma suaktyvinant tinklą, siekiant surinkti IP srautų srautus ir įdiegti anomalijų aptikimo sistemas, pagrįstas tinklo elgsenos analize (tinklo srauto stebėjimu).



2.4 pav. „NetFlow V9“ tipo surenkamų paketų pavyzdys [21]

„NetFlow“ yra „Cisco“ sukurtas tinklo protokolas, skirtas rinkti ir stebėti tinklo srauto duomenis, kuriuos generuoja „NetFlow“ palaikantys maršrutizatoriai ir komutatoriai. „NetFlow“ idėja buvo ta, kad pirmasis srauto paketas sukurs „NetFlow“ perjungimo įrašą jungiklyje ar maršrutizatoriuje, o vėliau šis įrašas bus naudojamas visiems vėlesniems to paties srauto paketams iki srauto pabaigos. Srautuose nėra faktinių paketinių duomenų, o komunikacijos metaduomenys. [22]

Kad būtų galima gauti tinklo srauto informaciją, „NetFlow V9“ duomenų perdavimo kadruose yra daug informacijos. „NetFlow“ paketo duomenų kadre esančiame lauko tipe yra tokia informacija (bet tuo neapsiribojama) (žr. 2.4 pav.):

Turėdamas šią informaciją, „Netflow“ leidžia įrenginiams sukurti kiekvieno srauto įrašą. Įrenginys siunčia šiuos įrašus „Netflow“ rinkėjui, kuris analizuoja duomenis ir teikia statistinius duomenis.

Šaltinyje [23] aprašomos savybės, kurios skirtos aptikti atakoms.

2.4.2 Anomalijų aptikimo modelis naudojant „Cisco Netflow“

Turint duomenų srautą galime nustatyti, kas yra tikėtinas normalus tinklo elgesys. Tam atlikti reikia žinoti, kokie parametrai, kurie yra išvadinti ankstesniame skyriuje (Apie „Cisco Netflow“), yra dabartiniame sraute. mus tarp dabartinio ir išsaugoto normalaus elgesio parametrų, sistema turi pranešti, kad galimai vyksta įtartini tinklo srauto veiksmai.

2.4.3 Atakų tipų nustatymas „Cisco Netflow“ sraute

„Cisco Netflow“ tipo srautas yra savybių iš įvairių paketų tipų (TCP, UDP, RDP ir kt.) reikšmių bendras rinkinys, pridedant papildomas reikšmes, kurios padeda nustatyti atakų tipą.

Šaltinyje [23] aprašoma tam tikroms tinklo atakoms atpažinti reikalingos savybės matomos lentelėje 2.1 lentelė. Joje išvadinamos savybės kurios priskiriamos tam tikriems atributams.

2.1 lentelė. Išskiriamos reikalingos savybės aptikti atakas „Cisco Netflow“ rinkinyje

Nr.	Duomenų rinkinio savybė	Priskiriamas atributui	Aprašas
1	icmp_dst_ip_b	icmp_smf	Tinklo transliavimas užliejamas ICMP paketais
2	icmp_src_ip	icmp_f	ICMP paketais užtvindomas taikinys
3	udp_dst_p	udp_f	Bandoma nutraukti sujungimą naudojant UDP srautą
4	tcp_f_s	tcp_syn_f	Užtvindymas SYN paketais
5	tcp_f_n_a	tcp_syn_f	Užtvindymas SYN paketais
6	tcp_f_n_f	tcp_syn_f	Užtvindymas SYN paketais
7	tcp_f_n_r	tcp_syn_f	Užtvindymas SYN paketais
8	tcp_f_n_p	tcp_syn_f	Užtvindymas SYN paketais
9	tcp_f_n_u	tcp_syn_f	Užtvindymas SYN paketais
10	tcp_dst_p	http_f	Bandoma nutraukti sujungimą naudojant HTTP srautą
11	tcp_src_dst_f_s	tcp_land	Pataikomas atakos tipas į bent kurį prievadą, naudojant SYN paketus
12	tcp_src_tftp	tcp_w32_w	Užtvindoma TFTP paslauga
13	tcp_src_kerb	tcp_w32_w	Užtvindoma Kerberos paslauga
14	tcp_src_rpc	tcp_w32_w	Užtvindoma RPC paslauga
15	tcp_dst_p_src	tcp_red_w	Naudoja HTTP serverio pažeidžiamumą
16	smtp_dst	smtp_b	Užtvindoma vieno kompiuterio SMTP jungtimis
17	udp_p_r_range	udp_reaper_w	Tikrinami prievandai: 80, 8080, 81, 88, 8081, 82, 83, 84, 1080, 3000, 3749, 8001, 8060, 8090, 8443, 8880 ir 10 000.
18	p_range_dst	tcp_udp_win_p	Keli preivadai, vienas adresas; vienas prievadas, keli adresai, Prievadai NBT, Samba, MS-SQL-S, VNC, RDP, 2222
19	udp_src_p_0	udp_0	Bandoma nutraukti paslaugą naudojant suskaidytą UDP srautą

Toliau kiekviena ataka yra aptinkama stebint tam tikrus atributus. Taip atakos yra susiejamos su stebimais atributais. Tai galime pamatyti 2.2 lentelė.

2.2 lentelė. Atakų tipų sąryšis su atributais „Cisco Netflow“ rinkinyje

Nr.	Atakos tipas	Stebimas atributas	Aprašas
1	„Smurf“	icmp_smf	Transliuojamos užklausos į tinklą aukos kompiuterio vardu
2	ICMP užtvindymas	icmp_f	Didelis ICMP paketų srautas
3	UDP užtvindymas	udp_f	Didelis srautas į DNS
4	TCP-SYN užtvindymas	tcp_syn_f	Didelis TCP paketų srautas su SYN ataka
5	HTTP užtvindymas	http_f	Didelis HTTP protokolų srautas
6	„LAND „	tcp_land	Taikinio IP adresas nurodomas tokio IP paketo antraštėje kaip paskirties ir išvykimo adresai, o bet koks atviras atakuojamos sistemos prievadas nurodomas kaip paskirties ir išvykimo prievadai
7	W32.Blaster kirminas	tcp_w32_w	Didelis srautas nuotolinio procedūrų iškvietimo (RPC) prievadų, skirtas TFTP, „Kerberos“ autentifikavimo prievadams
8	„Code Red“ kirminas	tcp_red_w	Naudoja žiniatinklio serverio pažeidžiamumą
9	„SAPM bots“	smtp_b	Pernelyg didelis SMTP ryšių skaičius
10	„Reaper“ kirminas	udp_reaper_w	„Reaper“ yra tinklas, kuris naudoja HTTP pagrįstus žinomų daiktų interneto pažeidžiamumų išnaudojimus
11	Skanavimo ataka	tcp_udp_win_p	Nenormalus ryšių skaičius iš vieno kompiuterio į vieną ar daugiau kitų kompiuterių
12	Paketų fragmentavimas	udp_0	Paslaugų atsisakymo atakos yra pagrįstos daugelio suskaidytų paketų naudojimu

Taip sujungtos atakos yra lengviau pastebimos analizuojamajame sraute, kurios tikrina jau tik atributų reikšmes, o ne atskirų savybių. Šias savybes lyginsime su kitų duomenų rinkinių savybėmis.

2.5 Duomenų rinkinio pasirinkimas tyrimui atlikti

Šiame darbe bus naudojamas Kanados kibernetinio saugumo instituto sudarytas sintetinis NSL-KDD duomenų rinkinys [27]. Įrašų skaičius NSL-KDD mokymosi ir bandymų rinkiniuose yra priimtinas. Dėl šio pranašumo galima eksperimentuoti su visu rinkiniu, nereikia atsitiktine tvarka pasirinkti mažos porcijos. Taigi skirtingų tyrimų darbų vertinimo rezultatai bus nuoseklūs ir palyginami.

NSL-KDD duomenų rinkinys, palyginti su originaliu KDD duomenų rinkiniu, turi šiuos pranašumus:

- siūlomuose bandymo rinkiniuose nėra pasikartojančių įrašų, todėl mokymosi pasiekimai nėra šališki dėl metodų, kurie turi geresnius rezultatus aptinkant pasikartojančius įrašus.
- iš kiekvienos sunkumų lygio grupės pasirinktų įrašų skaičius yra atvirkščiai proporcingas įrašų procentinei daliai pradiniame KDD duomenų rinkinyje. Dėl to atskirų mašininio mokymosi metodų klasifikavimo rodikliai skiriasi platesniu diapazonu, todėl šis rinkinys veiksmingiau ir tiksliau įvertina skirtingus mokymosi metodus.
- įrašų skaičius mokymo ir bandymų rinkiniai yra pagrįsti. Tas leidžia eksperimentuoti su pilnu duomenų rinkiniu, kuriam nereikia atsitiktinai pasirinkti mažos porcijos duomenų.

KDD mokymosi rinkinio nereikalingų įrašų statistika matoma žemiau pateiktoje lentelėje (žr. 2.3 lentelė.) .

2.3 lentelė. NSL-KDD duomenų rinkinių įrašų statistika [28]

Dataset	Number of Records:					
	Total	Normal	DoS	Probe	U2R	R2L
KDDTrain+20%	25192	13449 (53%)	9234 (37%)	2289 (9.16%)	11 (0.04%)	209 (0.8%)
KDDTrain+	125973	67343 (53%)	45927 (37%)	11656 (9.11%)	52 (0.04%)	995 (0.85%)
KDDTest+	22544	9711 (43%)	7458 (33%)	2421 (11%)	200 (0.9%)	2654 (12.1%)

Žinant visas šias įrašų kiekų reikšmes ir tai kad šio rinkinio autoriai teigia, jog visi tirti klasifikatoriai atpažino bent 86% testavimo rinkinio įrašų, galime lengvai įvertinti metodo efektyvumą.

2.6 NSL-KDD tinklo srauto struktūra ir atakų aptikimas jame

2.6.1 NSL-KDD tinklo srauto rinkinyje pateiktos savybės

Toliau norint atlikti veiksmams reikia suprasti duomenų rinkinio struktūrą ir kiekvieno stulpelio reikšmes.

2.4 lentelė. NSL-KDD duomenų rinkinyje pateiktos savybės

Savybė	**Savybės pavadinimas**	**Savybė**	**Savybės pavadinimas**	**Savybė**	**Savybės pavadinimas**
1	Ryšio trukmė	15	1, jei bandoma „su root“ komanda; 0 kitaip	29	% jungčių prie tos pačios paslaugos
2	Prisijungimo protokolas (pvz., TCP,UDP, ICMP)	16	„Root“ prieigų skaičius	30	% prisijungimų prie skirtingų paslaugų
3	Paskirties paslauga	17	Failų kūrimo operacijų skaičius	31	% jungčių su skirtingais kompiuteriais
4	Ryšio būsenos vėliava	18	„shell“ operacijų skaičius	32	Ryšių, turinčių tą patį paskirties pagrindinį kompiuterį, skaičius
5	Baitai, išsiųsti iš šaltinio į paskirties vietą	19	Prieigos kontrolės bylų operacijų skaičius	33	Ryšių, turinčių tą patį paskirties pagrindinį kompiuterį ir naudojančius tą pačią paslaugą, skaičius
6	Baitai, išsiųsti iš paskirties vietos į šaltinį	20	Siunčiamų komandų skaičius FTP sesijoje	34	% jungčių, turinčių tą patį paskirties pagrindinį kompiuterį ir

					naudojančios tą pačią paslaugą
7	1, jei ryšys yra iš / į tą patį pagrindinį kompiuterį / prievadą; 0 kitaip	21	1, jei prisijungimas priklauso „karštajam“ sąrašui; 0 kitaip	35	Įvairių dabartinio kompiuterio paslaugų procentas
8	Neteisingų fragmentų skaičius	22	1, jei prisijungimas yra „svečio“ prisijungimas; 0 kitaip	36	% jungčių su dabartiniu pagrindiniu kompiuteriu, turinčiu tą patį šaltinio prievadą
9	Skubių paketų skaičius	23	Ryšių su tuo pačiu pagrindiniu kompiuteriu, kaip ir dabartinis, skaičius per pastarąsias 2 sekundes	37	% prisijungimų prie tos pačios paslaugos, gaunamos iš skirtingų kompiuterių
10	„Karštų“ indikatorių skaičius	24	Prisijungimų su ta pačia paslauga, kaip dabartine, skaičius per pastarąsias dvi sekundes	38	% jungčių su dabartiniu kompiuteriu, kuriuose yra S0 klaida
11	Nepavykusių prisijungimų skaičius	25	% jungčių, kuriose yra „SYN“ klaidų	39	% ryšių su dabartiniu pagrindiniu kompiuteriu ir nurodyta paslauga, kuriuose yra S0 klaida
12	1, jei sėkmingai prisijungėte; 0 kitaip	26	% jungčių, kuriose yra „SYN“ klaidų	40	% jungčių su dabartiniu pagrindiniu kompiuteriu, kuriuose yra RST klaida
13	„Pavojingų“ sąlygų skaičius	27	% jungčių, turinčių REJ klaidų	41	% ryšių su dabartiniu pagrindiniu kompiuteriu ir nurodyta paslauga, kuriuose yra RST klaida
14	1, jei gautas administratoriaus teisės; 0 kitaip	28	% jungčių, turinčių REJ klaidų	42	Klasės etiketė

Kaip minėta ankstesniuose skyriuose, šiame tyrime mes naudojame NSL-KDD duomenų rinkinį. NSL-KDD yra patobulinta „KDD Cup 99“ duomenų rinkinio versija, kuri kenčia dėl daugybės nereikalingų įrašų. Į NSL-KDD duomenų rinkinyje pateiktos savybės parodytos lentelėje (žr. 2.4 lentelė.).

2.7 Atakų tipų nustatymas tinklo sraute

Jei neimtume jau įrašyto srauto, kaip apsimokymo pavyzdžio. Tada reikėtų realų tinklo srautą ir tikrinti TCP sluoksniuose vykstančias protokolų transakcijas. Transporto ir tinklo sąsajos sluoksniuose reikėtų analizuoti šaltinio ir gavėjo prievadus/bitus, turinio ilgio bitus, sekos numerius, visiems matomiems protokolams kaip: TCP,UDP,ICMP ir t.t. Toliau šie duomenys bus perduodami anomalijų aptikimo metodui kuris aprašomas kitame skyrelyje.

Kaip buvo rašyta anksčiau yra 4 pagrindiniai atakų tipai: DoS atakos, R2L atakos, U2R atakos ir zondavimo atakos. Toliau bus nagrinėjami požymiai pagal kuriuos bus galima aptikti šias atakas.

2.7.1 DoS atakų požymiai

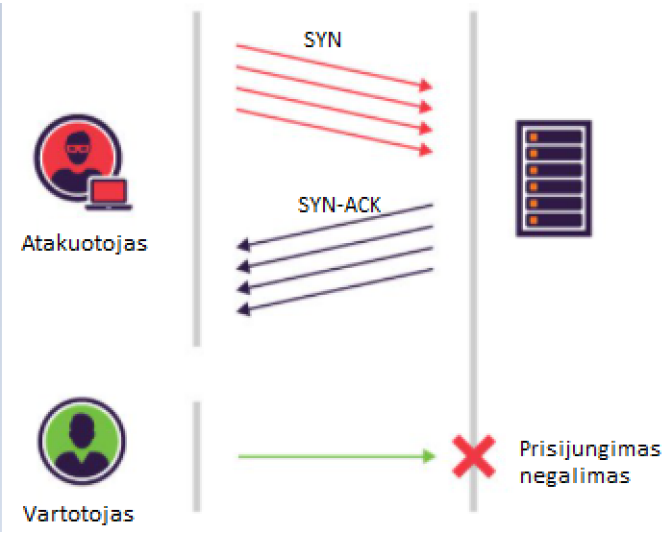
DoS atakos gali būti atliekamos daug būdų, tačiau populiariausi yra TCP SYN paketų užtvindymo ataka, HTTP paketų užtvindymas ir ICMP paketų užtvindymas. Visos šios atakos atliekamos naudojant skirtingus paketus, todėl reikia žiūrėti į kiekvieną paketų rūšį ir atskirti tai atakai būdingus požymius.

2.7.1.1 TCP SYN ataka

Pirmiausiai aptarkime TCP SYN tipo ataką Ją sudaro TCP paketų srautas su besikeičiančiomis vėliavėlėmis vykdant TCP jungtį kitaip vadinamą „rankų paspaudimą“ (žr. 2.5 pav.).

TCP paketas gali turėti 6 tipų vėliavėles:

1. SYN – užklausa sinchronizuoti sekos numerius;
2. ACK – patvirtinimo numeris yra veikiantis;
3. RST – panaikinti sujungimą;
4. FIN – duomenų siuntimo pabaiga;
5. URG – pirmumo rodyklės požymis;
6. PSH – prašoma, kad gavėjas perduotų proceso duomenis nelaukdamas, kol užsipildys buferis.



2.5 pav. TCP-SYN užtvindymo ataka

Šios atakos metu naudojamos 2 tipų vėliavėles: SYN ir SYN-ACK. Klientas išsiunčia didelį kiekį užklausų (TCP paketų su SYN vėliavėle) atakuojamam serveriui. Serveris savo ruožtu turi kiekvienam SYN paketui turi sugeneruoti SYN-ACK atsakymą, o jei šių užklausų ateina šimtais ar net tūkstančiais, tada sutrikdomas atakuojamos mašinos darbas ir ši mašina galimai bus nepasiekiama kol nebus apdorotos visos užklausos.

„Cisco Netflow“ sraute turime stebėti atributą „tcp_syn_f“.

2.7.1.2 HTTP paketų užtvindymo ataka

HTTP paketai naudojami kliento ir serverio komunikacijai užtikrinti. Tai vykdoma naudojant HTTP užklausas ir atsakymus. Dažniausiai naudojamos 2 tipų užklausos: GET ir POST.

HTTP GET ataka - naudojant šią atakos formą, keli kompiuteriai ar kiti įrenginiai yra sukonfigūruoti, kad iš tikslinio serverio būtų siunčiamos kelios vaizdų, failų ar kito turinio užklausos. Kai tikslą užplūs

gaunamos užklausos ir atsakymai, bus įvykdytas paslaugos trikdymo atakos tikslas ir turinio esančio šiame serveryje nubus galima pasiekti ir iš to norinčių klientų.

HTTP POST ataka - paprastai kai forma pateikiama svetainėje, serveris turi tvarkyti gaunamą užklausą ir perkelti duomenis į patvarumo sluoksnį, dažniausiai į duomenų bazę. Formos duomenų tvarkymo ir reikalingų duomenų bazės komandų vykdymo procesas yra gana intensyvus, palyginti su apdorojimo galios ir pralaidumo kiekiu, reikalingu POST užklausai siųsti. Ši ataka panaudoja išteklių sunaudojimo skirtumus serveryje, siunčiant daugybę pranešimų užklausų tiesiai į jį, kol jo resursai bus išnaudoti ir įvyks paslaugų sutrikdymo ataka.

„Cisco Netflow“ sraute turime stebėti „http_f“ atributą.

2.7.1.3 ICMP paketų užtvindymo ataka

+	Bitai: 0–7	8–15	16–31
0	Tipas	Kodas	Duomenų patikros suma

2.6 pav. ICMP paketo antraštė

Interneto valdymo pranešimų protokolo (ICMP) (žr. 2.6 pav.) užtvindymo ataka, dar vadinama „Ping“ užtvindymo ataka, yra DoS ataka, kurios metu užpuolikas bando išnaudoti tikslino įrenginio resursus siunčiant „ICMP *echo-request*“ paketus. Paprastai ICMP užklausos ir atsakymo pranešimai naudojami norint įsitikinti tinklo įrenginio veikimu, kad būtų galima nustatyti prietaiso būklę ir ryšį bei ryšį tarp siuntėjo ir įrenginio. Užtvindamas tikslą užklausų paketais, tinklas yra priverstas atsakyti vienodu atsakymų paketų skaičiumi. Dėl to taikinys tampa nepasiekiamas įprastam veikimui.

„Cisco Netflow“ sraute turime stebėti „icmp_f“ atributą.

2.7.2 R2L atakų požymiai

R2L ataka (angl. *Remote to Local*) yra ataka kurios metu atakuotojas bando perimti kito kompiuterio kontrolę neturėdamas šio kompiuterio prisijungimo duomenų.

R2L atakos yra vienos iš sunkiausiai aptinkamų, nes jos susijusios su tinklo lygio ir kompiuterio lygio funkcijomis ir požymiais. Todėl tikriname tinklo lygio požymius, tokius kaip ryšio trukmė ir reikalaujama paslauga, ir pagrindinio kompiuterio lygio požymius, pvz., nepavykusių prisijungimo bandymų skaičius.

NSL-KDD tinklo srauto duomenyse galime aptikti įvairių atakos tipų, kaip: „ftp write“, „gess pass“, „Imap“, „Multihope“, „phf“, „spy“, „warez“.

„Phf“ [24] ataka yra R2L ataka prieš žiniatinklio serverį, kuriame veikia „Phf CGI“ scenarijus. „Phf“ scenarijus turi pažeidžiamumą, kuris, pasinaudojus, leidžia nuotoliniams vartotojams vykdyti savavališkas žiniatinklio serverio komandas.

„Warezmaster“ (WM) ir „Warezclient“ išpuolis (WC) [25], kurie naudoja „anoniminio“ FTP pažeidžiamumus tiek „Linux“, tiek „Windows“.

2.7.3 U2R atakų požymiai

U2R ataka - tai ataka kai užpuolikas turi vietinę prieigą prie aukos mašinos ir bando įgyti super-vartotojo (administratoriaus) privilegijas.

U2R atakos apima semantines detales, kurias labai sunku užfiksuoti ankstyvoje stadijoje. Tokios atakos dažnai grindžiamos turiniu ir nukreiptos į programą. Taigi, naudojant U2R atakas, stebėjimui pasirinktos tokie požymiai kaip failų kūrinių skaičius ir iškviestų apvalkalo raginimų skaičius (angl. *number of shell prompts invoked*), o tokie požymiai kaip protokolas ir šaltinio baitai nepaisomi.

Galimi įvairūs atakos variantai: „buffer overflow“, „Rootkit“ ir kt.

Buferio perpildymas gali paveikti visų tipų programinę įrangą ar ne operacines sistemas. Paprastai jie atsiranda dėl netinkamai suformuotų įvestų arba nepakankamo vietos paskirstymo buferiui. Jei operacija perrašo vykdomąjį kodą, tai gali sukelti programos nenuspėjamą elgesį ir generuoti neteisingus rezultatus, prieigos prie atminties klaidas ar gedimus.

„Rootkit“ [26] atakos vykdymas gali būti automatizuotas arba užpuolikas gali jį įdiegti gavęs root arba administratoriaus prieigą. Šios prieigos gavimas yra tiesioginio užpuolimo prieš sistemą rezultatas, t. Y. Žinomo pažeidžiamumo (pvz., Privilegijų eskalavimo) arba slaptažodžio (gauto nulaužus ar socialinės inžinerijos taktiką, pvz., „Sukčiavimą“), panaudojimas. Įdiegus tampa įmanoma paslėpti įsilaužimą ir išlaikyti privilegijuotą prieigą. Visiškas sistemos valdymas reiškia, kad galima modifikuoti esamą programinę įrangą, įskaitant programinę įrangą, kuri kitu atveju gali būti naudojama jai aptikti ar apeiti.

Šioms atakai atpažinti reikia stebėti skubių paketų kiekį, stebėti ar gautos administratoriaus teisės, „shell“ komandų kiekį, ir. t.t.

2.7.4 Zondavimo atakų požymiai

Zondavimas yra ataka, kurios metu įsilaužėlis nuskaito mašiną ar tinklo įrenginį, kad būtų galima nustatyti trūkumus ir pažeidžiamumus, kurie vėliau gali būti panaudoti sukompromituoti sistemą ar mašiną. Todėl reikia stebėti sistemos prievadus ir tikrinti į tuos prievadus prisijungti norinčius mašinų adresus.

Ipsweep“ ataka yra stebėjimas, siekiant nustatyti, kurie kompiuteriai klausosi tinkle. Ši informacija yra naudinga užpuolikui rengiant išpuolius ir ieškant pažeidžiamų mašinų.

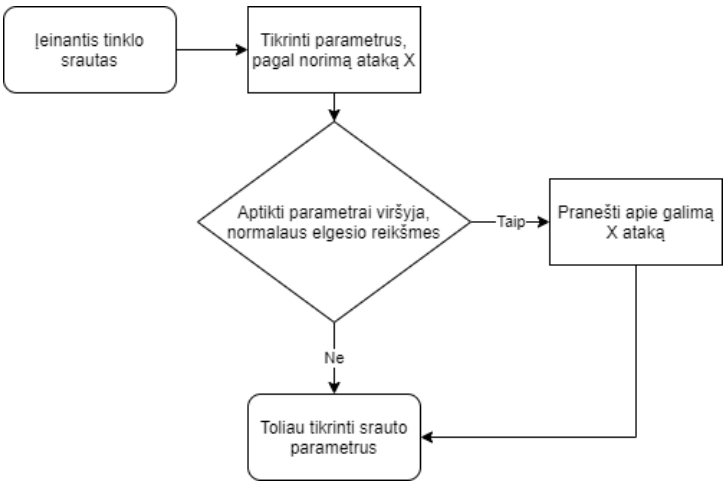
„Nmap“ yra bendros paskirties įrankis, skirtas atlikti tinklo nuskaitymus. „Nmap“ palaiko daugybę skirtingų tipų uosto nuskaitymo variantų: SYN, FIN ir ACK nuskaitymas tiek TCP, tiek UDP, taip pat ICMP (Ping) nuskaitymas. „Nmap“ programa taip pat leidžia vartotojui nurodyti, kuriuos uostus nuskaityti, kiek laiko laukti tarp kiekvienos prievado ir ar prievadus reikia nuskaityti nuosekliai, ar atsitiktine tvarka.

Kadangi, šios atakos yra panašios savo duomenų naudojimu, todėl jas galime stebėti lygiagrečiai naudojant tas pačias tinklo įrašo savybes.

Šio tipo atakos yra tokios: „Ipsweep“, „nmap“, „portsweep“. Paskutinės dvi atakos tikrina, atidarytus prievadus atakuojamoje mašinoje.

2.8 Taisyklių, skirtų atakoms tinkle aptikti, kūrimas

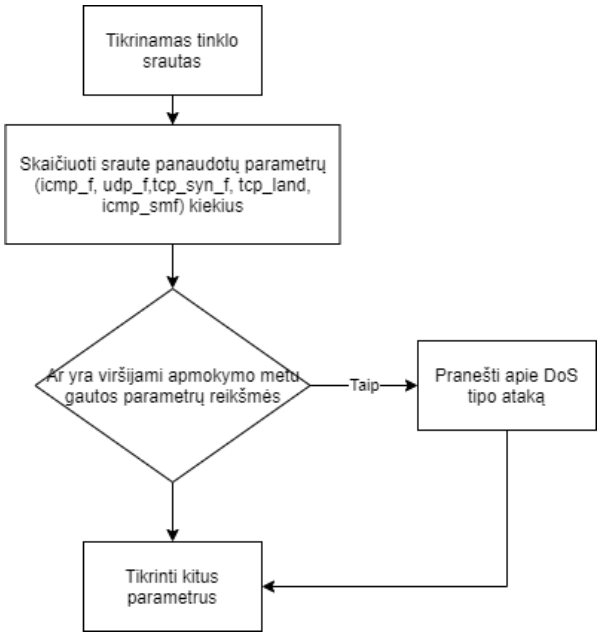
Išanalizavę atakas ir tinklo srautuose atskleistas savybes, galime prieiti prie klasifikatoriams skirtų taisyklių kūrimo. Šias taisykles reiktų kurti kiekvienam atakos tipui, pagal jam aptikti nurodytas savybes praeituose skyreliuose. Jas turime taikyti po apmokymo, o apmokymo metu reikia gauti reikšmes, kurias viršijus būtų galima teigti jog tai buvo ataka. Pavyzdžiui paimkime tinklo zondavimo ataką „nmap“. Ši ataka skenuoja visus atidarytus prievadus atakuojamoje mašinoje, todėl turime stebėti sujungimų skaičių ir atakos tikslą, jei dažnai naudojami to pačio adreso skirtingi prievadai, per tam tikrą laiko tarpą, galime teigti, jog vyksta panašūs veiksmai į zondavimo ataką. Bendru atveju atakos aptikimo taisyklę galima pavaizduoti taip (žr. 2.7 pav.):



2.7 pav. Taisyklės pavyzdys, nepriklausomai kokia tai ataka

Šis taisyklės tikrinimo modelis tinka visoms prieš tai aprašytoms atakoms, schemoje atakos pavadinimą atvaizduojame X simboliu, nes jau apmokintas anomalijų aptikimo metodas jau turi apskaičiuotas reikšmes, kurios reiškia galimą ataką, todėl jas reikia tik sulyginti X atakai skirtomis atpažinti savybėms, tokiu būdu nustatant atakos tipą ir jį pranešant tinklo administratoriui.

2.8.1 DoS tipo anomalijų aptikimo algoritmas naudojant „Cisco Netflow“

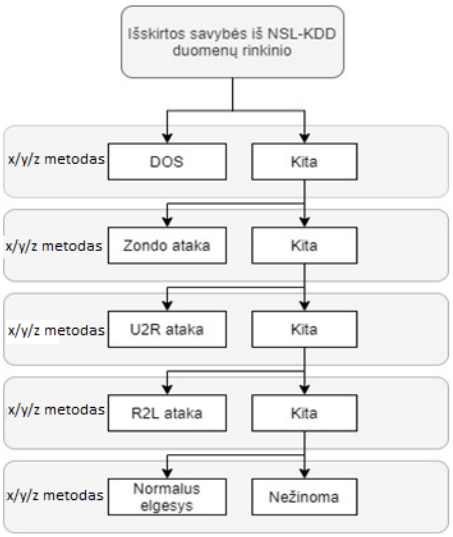


2.8 pav. DoS tipo atakos aptikimo, „Cisco Netflow“ sraute, filtro pavyzdys

Naudojantis panašiu modeliu į pavaizduotą skyrelyje 2.4.1. galime sumodeliuoti paprastą DoS atakų aptikimo modelį skirtą veikti su parametrais „Cisco Netflow“ tinkle. Pirmiausia reikia turėti jau apmokytus maršrutizatorius (sensorius) tinkle, kurie jau žino, koks yra normalus elgesys tinkle. Tada šie sensoriai skaičiuoja tinklo sraute praeinančius paketus ir jei jų skaičius yra didesnis nei kad normalaus elgesio tinklo sraute, tada privaloma išvesti pranešimą apie galimą ataką tinkle. Šį modelį galima pamatyti 2.8 pav.

2.9 Tinklo elgsenos anomalijų aptikimo modulio prototipas

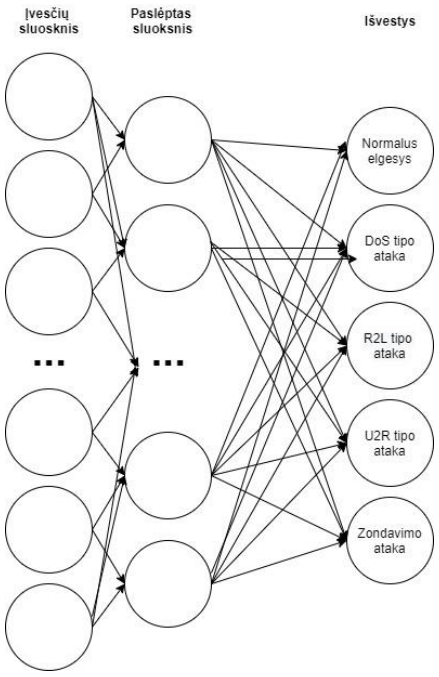
Anomalijų aptikimas vyks ankstesniuose skyreliuose aprašytu metodu, kurs nuskaitys NSL-KDD duomenų rinkinio duomenis, juos apdorojame ir sumažiname analizuojamų duomenų kiekį tolimesniam darbui, kurį atliks mašininio mokymosi algoritmas.



2.9 pav. Siūloma algoritmo architektūra

Mašininio mokymosi algoritmas realizuojamas pagrindžiant tyrimu, kuriame naudojamas penkių lygių klasifikavimo modelis, pagrįstas mašininio mokymosi algoritmų x/y/z deriniu. Architektūros struktūra pavaizduota šio skyrelio paveiksliuke (žr. 2.9 pav.).

Galutinis architektūros modelis bus parenkamas eksperimento metu įvertinus rezultatus.



2.10 pav. „Adam“ gilaus mokymosi metodo architektūra

Šis modulis bus lyginamas su „Adam“ gilaus mokymosi metodu, kuris bus sudarytas iš įvesčių sluoksnio, kuriame yra visos savybės, kelių paslėptų sluoksnių ir išvestis sluoksnio, kuriame jau išvedamas gautas rezultatas (žr. 2.10 pav.). „Adam“ metodas tinka dėl savo veiksmingumo su apdorojant iki 1 milijono siekiančius įrašų skaičius, bet to šio metodo apsimokymas vyksta labai greitai.

2.10 Projektavimo dalies išvados.

Identifikuotos tinklo srauto savybės, kurias galima gauti iš „Cisco Netflow“ įrašų ir panaudoti metodų apsimokymui. Tačiau negalime gauti R2L ir U2R atakų tipams skirtų savybių, todėl šios atakos nebus akcentuojamos.

Aprašyti tyrimui numatytų atakų specifiniai požymiai, pagal kuriuos bus galima atrinkti apsimokymui reikalingas srauto įrašų savybes.

Suprojektuotas apibendrintas eksperimentinio tyrimo algoritmas, kuris bus modifikuojamas tyrimo etape atsižvelgiant į gaunamus rezultatus.

Eksperimento metu gaunami rezultatai bus vertinami lyginant su „Adam“ gilaus mokymosi metodo rezultatais.

3 Mašininio mokymo metodų taikymo atakoms aptikti naudojant *Cisco Netflow* tinklo įrašus realizacija

3.1 Apsimokymo procesui reikalinga informacija apie esančias atakas.

Apsimokymo procese norime išmokyti atpažinti algoritmą atskirti DOS atakas, U2R bei R2L atakas ir zondavimo atakas. Aprašytas duomenų rinkinys turi atitinkamas atakas pagal jau išvardytas atakų klases (žr.

3.1 lentelė.):

3.1 lentelė. Atakų klasės KDD tinklo srauto įraše

Atakos klasė	**Atakos tipas**
DoS	„Smurf“, „Land“, „Pod“, „Teardrop“, „Neptune“, „Back“, ...
R2L	„ftp write“, „gess pass“, „Imap“, „Multihope“, „phf“, „spy“, „warez“, ...
U2R	„Perl“, „buffer overflow“, „Rootket“, ...
Zondavimo	„Ipsweep“, „nmap“, „portsweep“, ...

3.2 Atakų aptikimas NSL-KDD tinklo srauto rinkinyje

Norint efektyviai atpažinti šio tipo atakas reikia stebėti daug parametrų. Šaltinyje [29] aprašytos skirtingoms DoS atakoms atpažinti reikalingos savybės KDD tinklo surauto įraše. (žr. 3.2 lentelė.)

3.2 lentelė. DoS atakoms atpažinti reikalingos savybės iš KDD duomenų rinkinio

Atakos tipas	**Svarbiausios savybės**
„Smurf“	7
„Land“	2, 3, 5, 23, 24. 27, 28, 36, 40, 41
„Neptune“	4, 25, 26, 29, 30, 33, 34, 35, 38, 39
„Teardrop“	8
„Back“	10, 13

DOS tipo atakas atpažinti galima, NSL-KDD tinklo įrašuose stebint savybes: 2.3.4.5, kuriuos nusako tinklo protokolo naudojimą, paslaugos tipą, vėliavėlės reikšmę ir išsiųsto srauto dydį baitais.

R2L atakas galime aptikti stebint 12, 25, 27, 29, 33, 37, 38 savybes srauto įraše.

U2R atakų aptikimui, KDD tinklo surauto įraše reikia stebėti 9, 14, 18, 21, 33, 38 savybes.

Zondavimo atakas galime atpažinti stebint tokias savybes NSL-KDD tipo tinklo srauto įraše: 2, 3, 4, 5, 11, 12, 18, 19, 21 ir daugiau.

Visas stebimas savybes galime sudėlioti į lentelę. (žr. 3.3 lentelė.)

3.3 lentelė. DoS, R2L, U2R ir zondavimo atakų stebėjimui būtinos savybės

Atakos klasė	**Svarbiausios savybės**
DoS	11, 12, 23, 29, 31, 37

R2L	12, 25, 27, 29, 33, 37, 38
U2R	9, 14, 18, 21, 33, 38
Zondavimo	2, 3, 4, 5, 11, 12, 18, 19, 21, 22, 24, 27, 28, 31, 38, 40, 41

3.3 Tinklo elgsenos anomalijų aptikimo metodo modulis

3.3.1 Tinklo elgsenos anomalijų aptikimo modulio realizavimo priemonės

Norint sukurti tinklo anomalijų aptikimo metodą, reikia jį realizuoti programiškai. Tai darysime naudojant „Python“ programavimo kalbą kartu su „Visual Studio Code“ programavimo aplinka. Python kalba yra labai populiari, ji turi daug visiems pasiekiamų bibliotekų, kurios padeda atlikti veiksmus su neuroniniais tinklais ar kitais metodais, bei duomenų apdorojimu. Naudojama operacinė sistema „Windows 10“ leidžia naudotis visomis programinėmis kalbomis ir turi pritaikytas grafines sąsajas, įvairioms programavimo kalboms įgyvendinti.

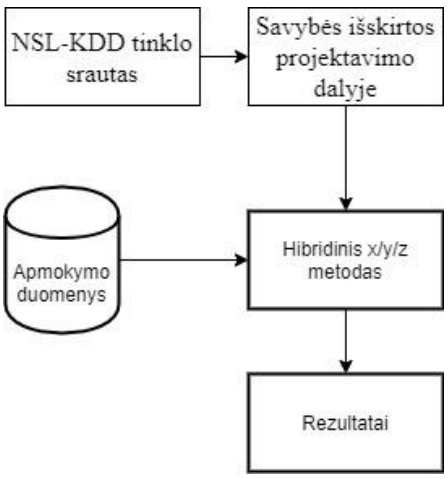
Projekto realizacijai panaudotos priemonės pateiktos lentelėje (žr. 3.4 lentelė.).

3.4 lentelė. Realizacijai naudojamos priemonės

Programavimo aplinka (IDE)	Visual Studio Code
Programavimo kalba	Python
Panaudotos bibliotekos	Sklearn, matplotlib, numpy, pandas, keras, tensorflow. datetime, os
Operacinė sistema	Windows 10
Aparatinė dalis	CPU:Intel(R) Core(TM) i7-3630QM CPU @ 2.40GHz Atmintis: 16,0 GB DDR3

3.4 Pasirinktos metodikos realizacija

3.4.1 Pasirinkto metodas aprašas



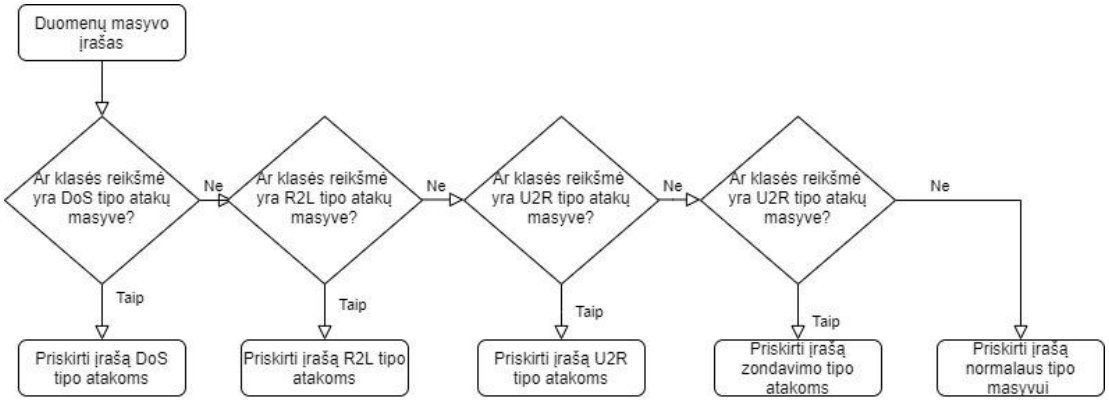
3.1 pav. Realizuojama tinklo anomalijų aptikimo sistemos architektūra

Apžvelgus visą prieš tai analizuotą medžiagą šaltinyje [6], nuspręsta pasirinkit hibridinį mašininio mokymosi metodą, kuris susidarys iš x, y ir z mašininio mokymosi metodų.

Šis metodas pirmiausia paims savybes iš NSL-KDD duomenų rinkinio. Tada jos bus paduodamos į 5 sluoksnių architektūros hibridinį mašininio mokymosi metodą, kuris yra parodytas 2.9 skyriuje, kurio nauda taip pat yra aprašyta tame pačiame skyrelyje.

3.5 Anomalijų aptikimo metodo realizacija

3.5.1 Duomenų rinkinio paruošimas



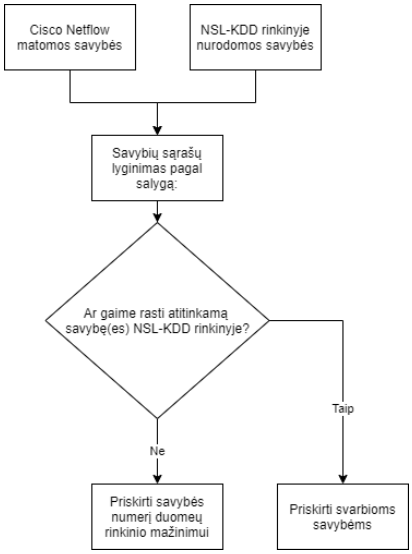
3.2 pav. Atakų priskyrimas atitinkamui atakos tipui

Duomenų rinkinio apdorojimas prisideda nuo atakų įrašų priskyrimo atakų klasėms, tai atliekame pirmiausiai priskiriant atitinkamas atakas į galimus 4 atakų tipus aprašytus projektavimo skyriuje. Taip pat, žinome ir šiems atakų tipams priklausančius atakų pavadinimus, todėl reikia surūšiuoti duomenų rinkinio įrašus pagal atakos rūšį, tai įgyvendiname atlikdami algoritmą pavaizduotą schemoje (žr.

3.2 pav.).

3.5.2 Duomenų rinkinio savybių mažinimas

Išanalizavus „Cisco Netflow“ paketo struktūrą, pastebime kad matome ne visas savybes, kokias matome NSL-KDD rinkinyje, todėl turime mažinti šį rinkinį kad sulygintume analizuojamas savybes, taip supaprastindami apsimokymo procesą. todėl turime atmesti daug savybių kurių negalime rasti „Cisco Netflow“ paketuose. Todėl esminių savybių atrinkimą galime pavaizduoti šiuo modeliu (žr. 3.3 pav.).



3.3 pav. Duomenų rinkinio savybių mažinimo modelis

Pagal projektavimo dalyje aprašytas ir išskirtas, pagal atakos tipą, savybes reikia apdoroti pradinį duomenų rinkinį. Šis procesas vykdomas iš duomenų rinkinio pašalinant parametrų stulpelius, tai atliekame įvykdę tokias kodo eilutes, kurios sumažina duomenų rinkinį iki reikiamo dydžio, į masyvą surašome nereikalingų stulpelių numerius. Atlikę šią operaciją gauname rezultatą, kuriame pirmas pavaizduotas pilnas duomenų rinkinys, o antroji dalis yra naujas dviejų dimensijų masyvas skirtas apmokymo metodui (žr. 3.4 pav.).

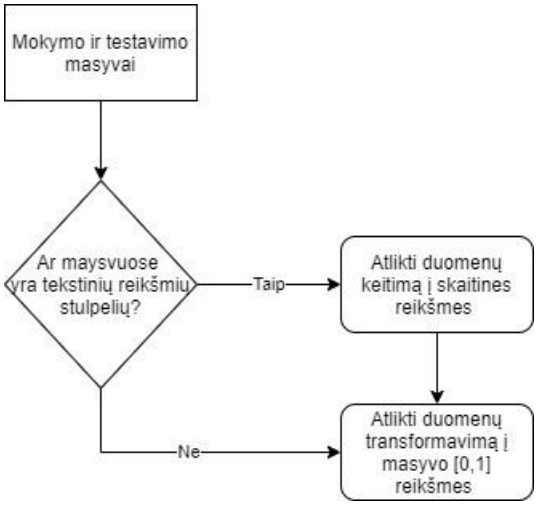
```
   duration protocol_type   service flag  src_bytes  dst_bytes  land  ...  dst_host_srv_diff_host_rate  dst_host_serror_rate  dst_host_srv_serror_rate  dst_host_rerror_rate  dst_host_srv_rerror_rate  outcome  difficulty
0         0           tcp  ftp_data   SF        491          0     0  ...                         0.00                  0.00                      0.00                  0.05                      0.00   normal          20
1         0           udp     other   SF        146          0     0  ...                         0.00                  0.00                      0.00                  0.00                      0.00   normal          15
2         0           tcp   private   S0          0          0     0  ...                         0.00                  1.00                      1.00                  0.00                      0.00  neptune          19
3         0           tcp      http   SF        232       8153     0  ...                         0.04                  0.03                      0.01                  0.00                      0.01   normal          21
4         0           tcp      http   SF        199        420     0  ...                         0.00                  0.00                      0.00                  0.00                      0.00   normal          21

[5 rows x 43 columns]

   duration protocol_type   service flag  src_bytes  urgent  num_failed_logins  ...  dst_host_srv_count  dst_host_srv_diff_host_rate  dst_host_serror_rate  dst_host_rerror_rate  dst_host_srv_rerror_rate  outcome  difficulty
0         0           tcp  ftp_data   SF        491       0                  0  ...                  25                         0.00                  0.00                  0.05                      0.00   normal          20
1         0           udp     other   SF        146       0                  0  ...                   1                         0.00                  0.00                  0.00                      0.00   normal          15
2         0           tcp   private   S0          0       0                  0  ...                  26                         0.00                  1.00                  0.00                      0.00  neptune          19
3         0           tcp      http   SF        232       0                  0  ...                 255                         0.04                  0.03                  0.00                      0.01   normal          21
4         0           tcp      http   SF        199       0                  0  ...                 255                         0.00                  0.00                  0.00                      0.00   normal          21

[5 rows x 27 columns]
```

3.4 pav. Sumažintų savybių duomenų rinkinys

Rezultate matyti, kad duomenų rinkinys tapo mažesnis, taip sutaupant laiko duomenų mokymosi metodui. Šie duomenys toliau perduodami apdorojimui.

3.5.3 Duomenų rinkinio savybių normalizavimas

Apdorojimo procese vykdomas duomenų normalizavimas remiantis „sklearn“ bibliotekoje esančiu „MinMaxScaler“ transformatoriumi, kurio metu visos vertinamos savybės yra paverčiamos į vienodas reikšmes, kurios yra masyvo [0,1] reikšmė. Tokiu būdu supaprastinamas duomenų pateikimas klasifikatoriaus metodui. Tačiau prieš šį veiksmą turime visas savybių reikšmes paversti į skaitines reikšmes, nes rinkinyje yra ir skaitinių ir tekstinių reikšmių. Neatlikus šio veiksmo, negalime įvykdyti reikšmių transformavimo naudojant „MinMaxScaler“. Tai galime atlikti įvykdę šį programinį kodą, kuriame vykdomas 3 stulpelių reikšmių transformavimas į skaitines reikšmes ir kitų stulpelių reikšmių normalizavimas rėžio [0,1] reikšmes (žr. 3.5 pav.).



3.5 pav. Simboliais aprašytų savybių keitimas skaitinėmis

Atlikus šiuos veiksmus galime pateikti duomenų rinkinį pateikti toliau klasifikatoriaus apmokymui ir testavimui.

3.5.4 Realizuojamų metodų aprašai

3.5.4.1 Sudėtinio klasifikatoriaus metodo aprašas

Toliau bus naudojamas sudėtinis mašininio mokymosi metodas, kuris susidaro iš trijų mašininio mokymosi metodų: KNN klasifikatoriaus ir atsitiktinių miškų klasifikatoriaus. Pirmiausiai, turime apsirašyti sudėtinį klasifikatorių sudarančius klasifikatorius masyve. Jie yra aprašomi tokiomis komandomis (žr. 3.6 pav.)

```
estimators = [
 ('knn', KNeighborsClassifier(n_neighbors=7)),
 ('rf' , make_pipeline(StandardScaler(), RandomForestClassifier(criterion='entropy', max_depth=30, n_estimators=48, random_state=0)))
]
```

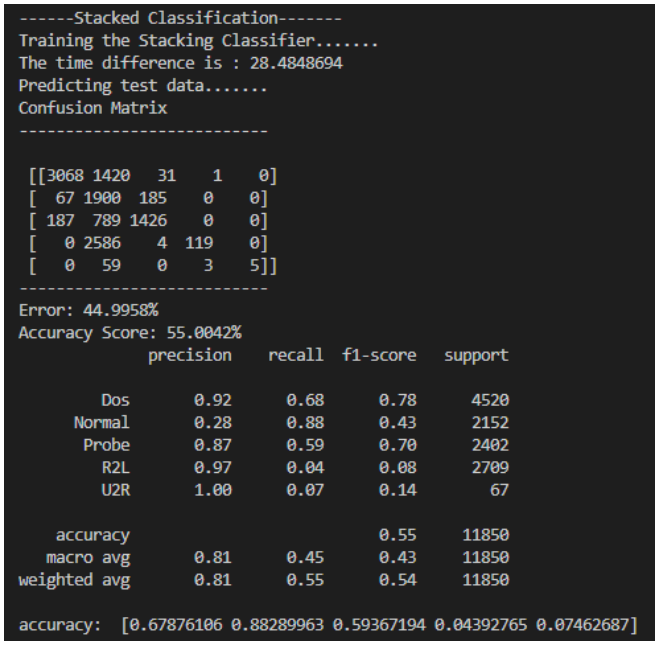
3.6 pav. Klasifikatorių aprašai, naudojami apjungimui

Šie du klasifikatoriai sujungiami naudojant “sudėtinį klasifikatorius” (angl. *Stacked classifiers*) ir jis prideda papildomą klasifikatorių, šiuo atveju naudojamas logistinės regresijos klasifikatorius. (žr. 3.7 pav.)

```
clf = StackingClassifier(estimators=estimators,n_jobs=-1, final_estimator=LogisticRegression())
```

3.7 pav. Sudėtinio klasifikatoriaus aprašas

Šis sudėtinis klasifikatorius susidarys iš prieš tai jau nurodytų klasifikatorių juos sujungiant į vieną.



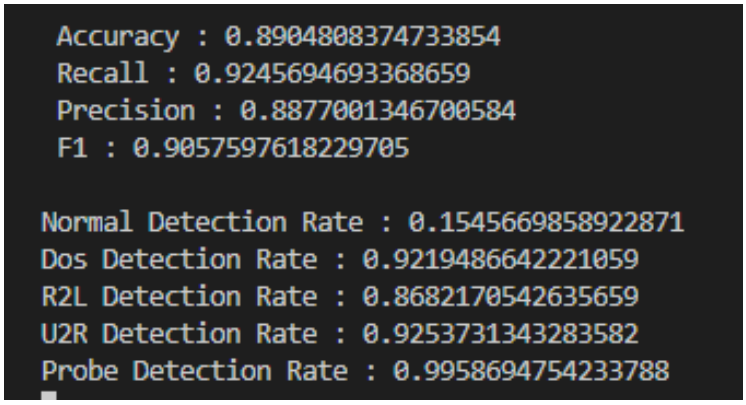
3.8 pav. Sudėtinio klasifikatoriaus prototipo rezultatai

Įvykdžius programinį kodą gauname rezultatus (žr. 3.8 pav.), kurie mums padės sulyginti šį metodą su kitais.

3.5.4.2 Gilaus mokymosi metodo „Adam“ aprašas

Sudėtinio klasifikatoriaus metodas bus lyginamas su „Adam“ gilaus mokymosi metodu, todėl turime sudaryti veikiantį modelį, kuris bus priklausomas nuo pateiktų duomenų. Metodas bus apmokomas 20 epochų, apmokymo duomenys yra paskirstyti, kad 20% mokymo duomenų, kurie bus naudojami kaip patvirtinimo duomenys.

Norint pasirengti eksperimentinei daliai pasirinktas modelis privalo išvesti rezultatus, kurie atskleistų modelio efektyvumą, tai galime padaryti pateikiant anomalijų aptikimo tikslumą ir kitas reikšmes. (žr. 3.9 pav.)



3.9 pav. Anomalijų aptikimo modelio realizacijos rezultatai, pateikiant atakų aptikimo tikslumą (naudojant „Adam“ gilaus mokymosi metodą)

Matant gautus rezultatus, žinome kad modelis gali būti dar efektyvesnis, todėl reikia keisti modelio apmokymui skirtus parametrus norint gauti rezultatą, atitinkantį apsibrėžtuose tiksluose.

3.6 Realizuoto modulio ypatumai

Realizuotas modulis yra labai lengvai keičiamas, modelio atžvilgiu. Galime lengvai manipuliuoti kokias savybes iš rinkinio galime analizuoti. Tačiau šis modulis yra pritaikytas tik tam tikros struktūros duomenims, t.y. NSL-KDD varžyboms sudaryta duomenų struktūra, todėl realizuojant šį modulį kitais atvejais, bus reikalingi dideli pakeitimai duomenų struktūros viduje ir duomenų išgavime (ne visada pavyks išgauti tiek daug duomenų). Kitas privalumas kad galime tirti atskirus klasifikatorius ir jų junginius. Šiuo atveju, nedidelėmis pastangomis, galime pakeisti kiekvieno klasifikatoriaus vidinius parametrus pakeitę vos kelias reikšmes.

3.7 Realizacijos išvados

Eksperimentiniam tyrimui pasirinktas Kanados kibernetinio saugumo instituto NSL-KDD duomenų rinkinys, kuris gerai atitinka tyrimo uždavinių sąlygas ir turi pakankamą reikalingų savybių rinkinį.

Ištyrus šį rinkinį atrinktos DoS, R2L, U2R ir zondavimo atakų aptikimui būtinos savybės.

Programiškai realizuotas tinklo elgsenos anomalijų aptikimo modulis, kuriame galima kaitalioti parametrus siekiant geriausių apsimokymo rezultatų.

Eksperimentavimo eigoje surastos ir realizuotos duomenų rinkinio sumažinimo ir normalizavimo procedūros, kurios padeda išgauti geresnius apsimokymo rezultatus.

4 Mašininio mokymo metodų taikymo atakoms aptikti naudojant *Cisco Netflow* tinklo įrašus tyrimas

4.1 Tyrimo tipas

Šiame darbe bus naudojamas taikomasis tyrimo tipas, kuris šiame darbe taikomas kaip viešai pasiekiamų duomenų rinkinio panaudojimas norint įvertinti mašininio mokymosi metodų pritaikymą realiose situacijose.

4.2 Tyrimo metodika

Tyrimo metu bus įvertinami analitinėje dalyje aprašyti mašininio mokymosi metodai:

- įvairūs klasifikatoriai (Atsitiktiniai miškai, KNN, SVM, Logistinės regresijos, SGD ir MLP klasifikatoriai);
- ir geriausiai pasirodžiusių kombinacija;
- gilaus mokymosi algoritmas naudojantis „Adam“ optimizatorių.

Tyrimui naudojamas duomenų rinkinys apima normalaus tinklo elgesio ir 4 atakos tipų (DoS, R2L, U2R ir zondavimo) įrašus. Kiekviename įraše yra 42 savybių reikšmės. Pilną rinkinį sudaro 125 tūkstančiai apsimokymui skirtų įrašų ir 22 tūkstančiai įrašų skirtų testavimui, o sumažinto iki 20% duomenų mokymo rinkinį sudaro 25 tūkstančiai įrašų ir 11 tūkstančių įrašų testavimui.

Tai sugrupuota į 4 scenarijus:

- VV : visas duomenų rinkinys ir visos savybės;
- VA : visas duomenų rinkinys ir atrinktos savybės;
- 20V: 20% apsimokymo duomenų ir visos savybės;
- 20A: 20% apsimokymo duomenų ir atrinktos savybės.

Kitame skyrelyje aprašyti atakų aptikimo metodų efektyvumo vertinimo kriterijai, kuriais remsimės darydami vertinimo išvadas.

4.3 Kriterijai metodų efektyvumui įvertinti

NBAD vertinimo kriterijai priklauso nuo painiavos matricos įvertinimo kaip klasifikavimo problemos. Painiavos matricos tikslas yra palyginti tikras ir numatomas etiketes. Pripažįstama, kad įsilaužimo aptikimo problemą sudaro dvi klasės: normalioji ir išpuolių. Toliau pateikta informacija buvo aptarta [5] dokumente.

Sąvokos TP (tikrasis teigiamas) ir TN (tikrasis neigiamas) žymi teisingai numatytas sąlygas, o FP (klaidingai teigiamas) ir FN (klaidingai neigiamas) neteisingai klasifikuojamas. TP ir TN nurodo atitinkamai klasifikuotus išpuolių ir normalius įrašus, ir, atvirkščiai, FP ir FN nurodo atitinkamai neklasifikuotus normalius ir išpuolių įrašus. Šie keturi terminai naudojami kuriant toliau sekančias IDS vertinimo priemones ir sumaišymo matricą.

- **Tikslumas (A)** (angl. *accuracy*) yra metrika, apskaičiuojanti bendrą IDS modelio nustatytų aptikimo ir melagingų aliarmų procentą, kuris atspindi bendrą bet kurios IDS sėkmės procentą ir yra apskaičiuojamas:

$$A = \frac{(TN + TP)}{(TP + FP + TN + FN)} \qquad (2)$$

- **Aptikimo koeficientas (R)** (angl. *recall*), dar vadinamas tikru teigiamu koeficientu (TPR) arba jautrumu, yra teisingai klasifikuotų kenksmingų atvejų dalis nuo viso kenksmingų vektorių skaičiaus ir yra apskaičiuojamas kaip:

$$R = \frac{TP}{(FP + TP)} \quad (3)$$

- **Preciziškumas (P)** (angl. *precision*) klasei nusako kiek iš visų numatytų klasės verčių, pvz., 1, kiek iš tikrųjų priklauso 1 klasei. Preciziškumas apskaičiuojamas kaip :

$$P = \frac{TP}{(FN + TP)} \quad (4)$$

- **F1 reikšmė (F1)** naudojama norėdami turėti sujungtą tikslumo ir aptikimo koeficiento efektą. F1 reikšmė yra tikslumo ir preciziškumo harmoninis vidurkis:

$$F1 = \frac{2}{(\frac{1}{P} + \frac{1}{R})} \quad (5)$$

Eksperimentais nustatytas metodų efektyvumas atskiroms atakų klasėms toliau pateiktas lentelėse/ar diagramose pagal tris metrikas: preciziškumas, aptikimo koeficientas ir F1 reikšmė.

4.4 Pavienių klasifikatorių veikimo įvertinimas pagal tyrimo scenarijus

4.3 skyriuje aprašėme kokiomis savybėmis vertinsime metodų efektyvumą, todėl šiame skyriuje jas panaudosime atvaizduojant tyrimo rezultatus.

Toliau bus tiriami scenarijai aprašyti 4.2 skyriuje.

4.4.1 VV scenarijus (visas duomenų rinkinys ir visos savybės)

Šiame skyriuje tiriame pilną duomenų rinkinį, kurį sudaro 125 tūkstančiai apsimokymui skirtų įrašų ir 22 tūkstančiai įrašų skirtų testavimui. Taip pat šiame duomenų rinkinyje tirsime jau ne visas savybes, o tik tam tikras savybes, išskirtas tyrimo skyriuje.

Metodų efektyvumas atskiroms atakų klasėms galimas panaudojant tris metrikas, tai preciziškumas, aptikimo koeficientas ir F1 reikšmė, toliau bus atvaizduoti paveiksliukai atitinkantys kiekvieną metodą.

	precision	recall	f1-score
Dos	0.96	0.77	0.86
Normal	0.65	0.97	0.78
Probe	0.85	0.59	0.70
R2L	0.98	0.08	0.15
U2R	0.40	0.03	0.06

4.1 pav. Atsitiktinių miškų metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VV

Atsitiktinių miškų klasifikatorius naudojant visas savybes ir pilną duomenų rinkinį, labai gerai aptinka DoS ir R2L tipų atakas, tačiau prastai aptinka U2R atakos tipo įrašus (žr. 4.1 pav.).

	precision	recall	f1-score
Dos	0.94	0.80	0.86
Normal	0.67	0.93	0.78
Probe	0.71	0.67	0.69
R2L	0.69	0.04	0.07
U2R	0.74	0.25	0.38

4.2 pav. KNN metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VV

KNN klasifikatorius geriausiai aptinka DoS tipo atkas ir vidutiniškai (apie 70 procentų) aptinka normalaus elgesio ir zondavimo atakos įrašus (žr. 4.2 pav.).

	precision	recall	f1-score
Dos	0.95	0.80	0.87
Normal	0.66	0.93	0.77
Probe	0.71	0.65	0.68
R2L	0.97	0.10	0.17
U2R	0.74	0.37	0.50

4.3 pav. SVM metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VV

SVM klasifikatorius veikimas labai panašus į KNN klasifikatoriaus, tačiau šis klasifikatorius prasčiau atpažino R2L ir U2R tipo atakas (žr. 4.3 pav.).

	precision	recall	f1-score
Dos	0.97	0.79	0.87
Normal	0.66	0.93	0.77
Probe	0.74	0.74	0.74
R2L	0.61	0.05	0.09
U2R	0.88	0.22	0.36

4.4 pav. Logistinės regresijos metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VV

Logistinės regresijos klasifikatorius labai gerai aptinka DoS tipo atakas, tačiau prasčiau aptinka R2L ir U2R tipo atakas (žr. 4.4 pav.).

	precision	recall	f1-score
Dos	0.96	0.73	0.83
Normal	0.61	0.93	0.74
Probe	0.68	0.57	0.62
R2L	0.00	0.00	0.00
U2R	0.00	0.00	0.00

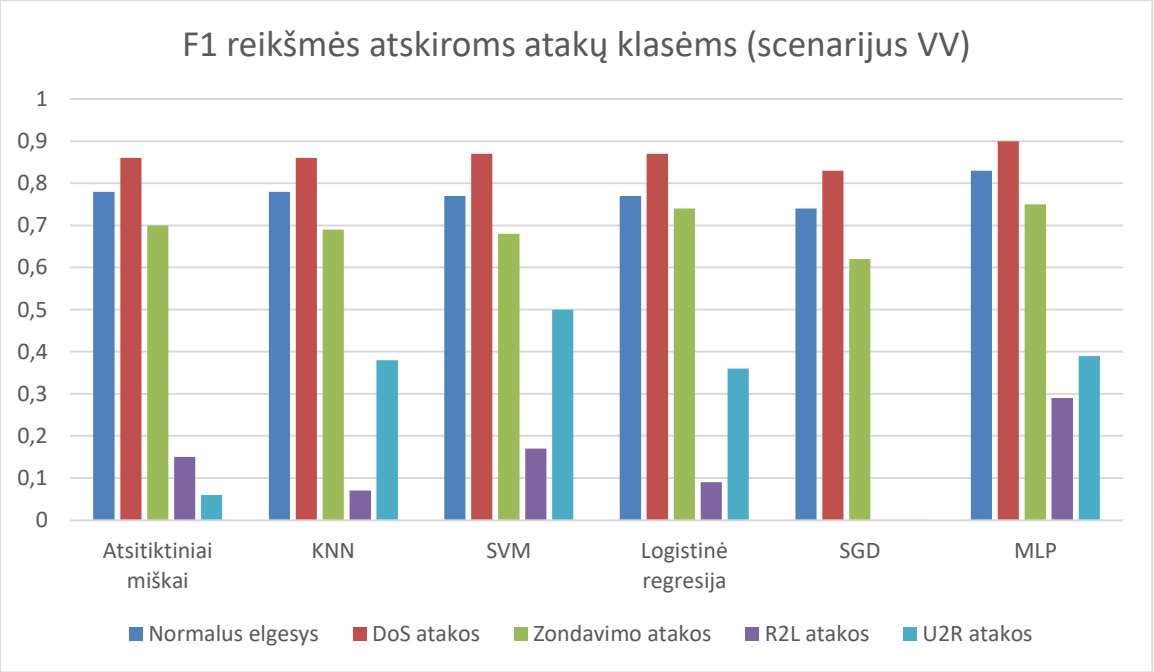
4.5 pav. SGD metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VV

SGD tipo klasifikatorius visiškai netinka R2L ir U2R tipų atakų atpažinimui, nes nebuvo nei vieno teisingai klasifikuoto įrašo šioms klasėms (žr. 4.5 pav.).

	precision	recall	f1-score
Dos	0.96	0.84	0.90
Normal	0.73	0.97	0.83
Probe	0.86	0.66	0.75
R2L	0.53	0.20	0.29
U2R	0.63	0.28	0.39

4.6 pav. MLP metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VV

MLP klasifikatorius yra geriausiai įvertintas iš visų anksčiau tirtų klasifikatorių, nes jo F1 reikšmė yra didžiausia (žr. 4.6 pav.).

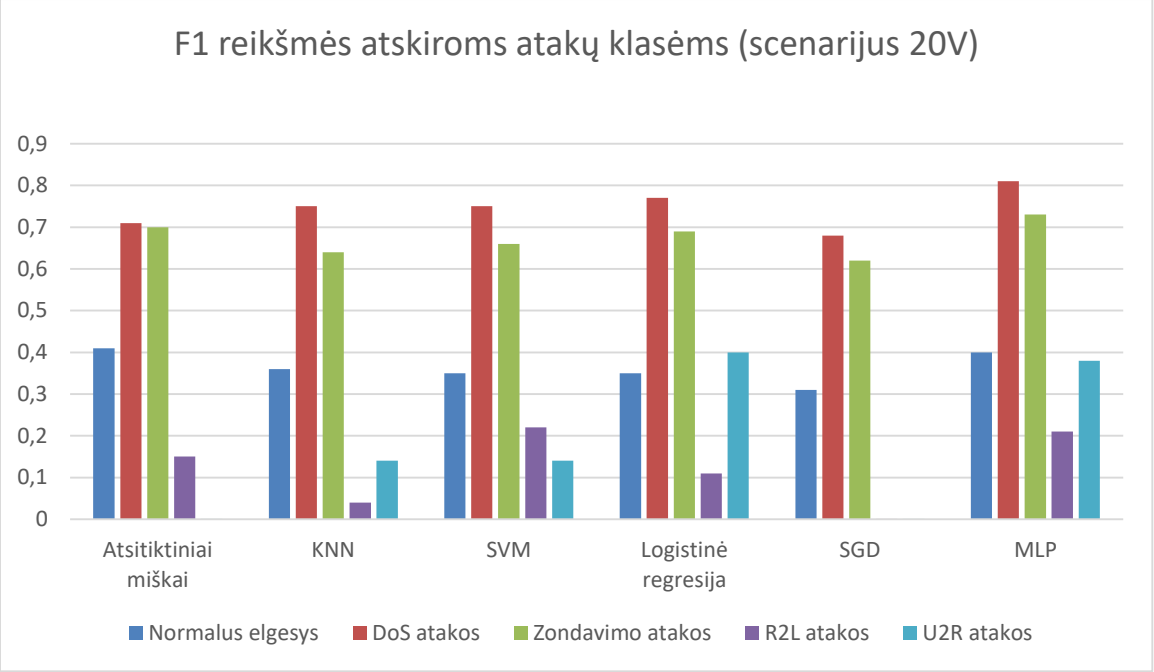


4.7 pav. Klasifikatorių išvestos F1 reikšmių, įvykdžius scenarijų VV, diagrama

Iš gautų bendrų rezultatų lentelėje (žr. 7.1 lentelė. ir 4.7 pav.) ir rezultatų paveiksluose, galime pamatyti kad su daug duomenų geriausiai veikiantis klasifikatorius yra MLP, nors logistinės regresijos klasifikatorius aukštu tikslumu aptinka normalų elgesį, tačiau jis visai neatpažino R2L ir U2R atakų. Kitas gerai pasirodęs klasifikatorius yra logistinės regresijos metodas. Todėl galime svarstyti šiuos du klasifikatorius kaip pagrindinius kandidatus būti sudėtiniame klasifikatoriuje.

4.4.2 20V scenarijus (sumažintas duomenų rinkinys ir visos savybės)

Šiame tyrimo scenarijuje tiriame sumažintą duomenų rinkinį, kurį sudaro 25 tūkstančiai apsimokymui skirtų įrašų ir 11 tūkstančių įrašų skirtų testavimui. Taip pat šiame duomenų rinkinyje tirsime tas pačias savybes kaip ir ankstesniame tyrimo scenarijuje.



4.8 pav. Klasifikatorių išvestos F1 reikšmių, įvykdžius scenarijų 20V, diagrama

Akivaizdžiai matome, kad šis duomenų rinkinys pasirodo prasčiau, nei pilnas (žr. 7.2 lentelė. ir 4.8 pav.).

	precision	recall	f1-score
Dos	0.92	0.58	0.71
Normal	0.27	0.88	0.41
Probe	0.87	0.58	0.70
R2L	1.00	0.08	0.15
U2R	0.00	0.00	0.00

4.9 pav. Atsitiktinių miškų metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų 20V

Iš atsitiktinių miškų klasifikatoriaus rezultatų matome, kad šis metodas neaptinka U2R tipo atakos, su pakankamai mažu kiekiu duomenų (žr. 4.9 pav.).

	precision	recall	f1-score
Dos	0.85	0.67	0.75
Normal	0.24	0.68	0.36
Probe	0.69	0.60	0.64
R2L	0.55	0.02	0.04
U2R	0.83	0.07	0.14

4.10 pav. KNN metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų 20V

KNN metodas geriausiai aptinka DoS ir zondavimo tipo atakas, tačiau labai prastai aptinka U2R, R2L ir normalaus tipo įrašus (žr. 4.10 pav.).

	precision	recall	f1-score
Dos	0.86	0.67	0.75
Normal	0.24	0.68	0.35
Probe	0.75	0.60	0.66
R2L	0.97	0.12	0.22
U2R	1.00	0.07	0.14

4.11 pav. SVM metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų 20V

SVM klasifikatorius geriau aptiko R2L, U2R ir zondavimo tipo atakas, lyginat su KNN klasifikatoriumi (žr. 4.11 pav.).

	precision	recall	f1-score
Dos	0.94	0.66	0.77
Normal	0.24	0.67	0.35
Probe	0.72	0.67	0.69
R2L	0.54	0.06	0.11
U2R	0.75	0.27	0.40

4.12 pav. Logistinės regresijos metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų 20V

Logistinės regresijos metodas gerai aptinka tik DoS ir zondavimo tipų atakas (žr. 4.12 pav.).

	precision	recall	f1-score
Dos	0.91	0.54	0.68
Normal	0.20	0.68	0.31
Probe	0.68	0.56	0.62
R2L	0.00	0.00	0.00
U2R	0.00	0.00	0.00

4.13 pav. SGD metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų 20V

SGD klasifikatorius lyginant su kitais klasifikatoriais, pakankamai gerai aptiko DoS, zondavimo ir normalaus tipo įrašus, tačiau visiškai neaptinka R2L ir U2R tipo atakų (žr. 4.13 pav.).

	precision	recall	f1-score
Dos	0.93	0.72	0.81
Normal	0.29	0.67	0.40
Probe	0.72	0.74	0.73
R2L	0.45	0.14	0.21
U2R	0.74	0.25	0.38

4.14 pav. MLP metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų 20V

MLP klasifikatorius rodo geriausius rezultatus iš visų tirtų klasifikatorių, nes parodė geriausius rezultatus aptinkant visų tipų įrašus (žr. 4.14 pav.).

Iš gautų rezultatų galime matyti prastesnius klasifikavimo rezultatus nei ankstesniame scenarijuje, bet tai galime paaiškinti sumažėjusiu apmokymo duomenų skaičiumi. Tačiau tie patys klasifikatoriai pasirodė geriausiai ir tiriant šį duomenų rinkinį, todėl sudarant sudėtinį klasifikatorių galime svarstyti KNN, MLP ir SVM metodų naudojimą ir juos sujungti į vieną sudėtinį klasifikatorių.

4.4.3 VA scenarijus (visas duomenų rinkinys ir atrinktos savybės)

4.4.3.1 Pavienių klasifikatorių tyrimas

Šiame scenarijuje naudosime tuos pačius duomenų rinkinius, tačiau iš jų panaikinsime dalį savybių, kurios buvo aptartos analizės skyriuje ir tokiu būdu geriau įvertinsime klasifikatorių veiksmingumą su duomenimis atitinkančiais „Cisco Netflow“ tinklo įrašų duomenis.

Šiame scenarijuje naudojamas duomenų rinkinys atitinka pirmo tyrimo scenarijaus, tačiau yra sumažintas tiriamų savybių skaičius nuo 41 iki 25.

	precision	recall	f1-score
Dos	0.96	0.77	0.86
Normal	0.66	0.97	0.78
Probe	0.76	0.61	0.67
R2L	0.97	0.02	0.04
U2R	0.80	0.18	0.29

4.15 pav. Atsitiktinių miškų metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VA

Atsitiktinių miškų metodas labai gerai aptiko normalaus elgesio įrašus (net 97%), tačiau kaip ir ankstesniuose scenarijuose prastai aptinka R2L ir U2R tipo įrašus. (žr. 4.15 pav.).

	precision	recall	f1-score
Dos	0.92	0.78	0.85
Normal	0.67	0.92	0.78
Probe	0.66	0.65	0.65
R2L	0.72	0.09	0.16
U2R	0.68	0.34	0.46

4.16 pav. KNN metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VA

KNN klasifikatoriaus rezultatai labai nežymiai sumažėjo visuose įrašų klasėse, lyginant su pirmuoju tyrimo scenarijumi (žr. 4.16 pav.).

	precision	recall	f1-score
Dos	0.83	0.80	0.82
Normal	0.66	0.92	0.77
Probe	0.83	0.55	0.66
R2L	0.01	0.00	0.00
U2R	0.00	0.00	0.00

4.17 pav. SVM metodo rezultatai atskiroms atakų klasėms, vykdant scenarijų VA

SVM klasifikatoriaus rezultatai suprastėjo, lyginant su pirmuoju tyrimo scenarijumi, nes dabar šis metodas neaptinka R2L ir U2R tipų atakos įrašų (žr. 4.17 pav.).

	precision	recall	f1-score
Dos	0.88	0.78	0.83
Normal	0.65	0.92	0.77
Probe	0.82	0.61	0.70
R2L	0.06	0.00	0.01
U2R	0.63	0.18	0.28

4.18 pav. Logistinės regresijos metodo rezultatai atskiroms atakų klasėms, kai naudojamas sumažintas savybių skaičius pilname duomenų rinkinyje

Logistinės regresijos klasifikatoriaus rezultatai suprastėjo, nes dabar šis metodas beveik neaptinka R2L tipų atakos įrašų (žr. 4.18 pav.).

	precision	recall	f1-score
Dos	0.81	0.73	0.77
Normal	0.63	0.93	0.75
Probe	0.90	0.42	0.57
R2L	0.00	0.00	0.00
U2R	0.00	0.00	0.00

4.19 pav. SGD metodo rezultatai atskiroms atakų klasėms, kai naudojamas sumažintas savybių skaičius pilname duomenų rinkinyje

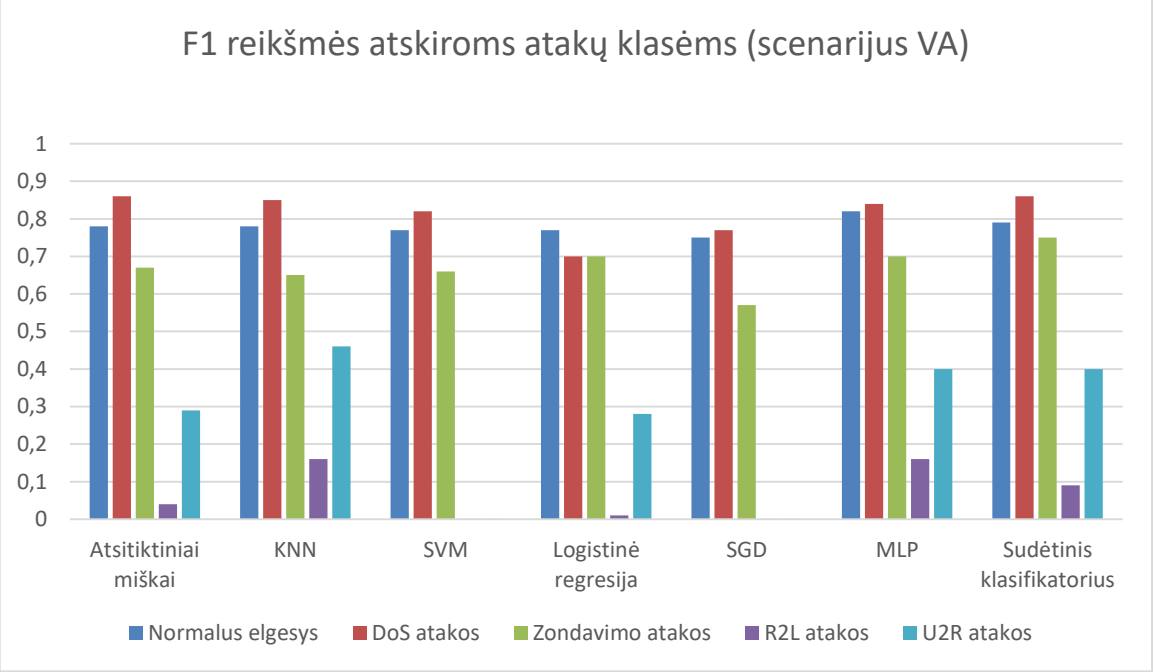
SGD klasifikatoriaus rezultatai nepakito, lyginant su pirmuoju tyrimo scenarijumi (žr. 4.19 pav.).

	precision	recall	f1-score
Dos	0.94	0.77	0.84
Normal	0.72	0.96	0.82
Probe	0.70	0.70	0.70
R2L	0.33	0.10	0.16
U2R	0.61	0.30	0.40

4.20 pav. MLP metodo rezultatai atskiroms atakų klasėms, kai naudojamas sumažintas savybių skaičius pilname duomenų rinkinyje

MLP klasifikatorius ir šiuo atveju pasirodė geriausiai iš visų tirtų klasifikatorių, visuose tiriamų įrašų klasėse matome didžiausias F1 reikšmes (žr. 4.20 pav.).

Ištyrę šiame scenarijuje gautus rezultatus, matome nežymų sumažėjimą visų tipų klasifikatorių rezultatuose, tačiau tai galime paaiškinti sumažėjusiu tiriamų savybių kiekiu.



4.21 pav. Klasifikatorių išvestos F1 reikšmių, įvykdžius scenarijų VA, diagrama

Išanalizavus gautus rezultatus (žr. 7.3 lentelė. ir 4.21 pav.) matome, kad geriausiai pasirodė MLP klasifikatorius (gauta F1 reikšmė 0,74), toliau rikiuojasi KNN ir atsitiktinių miškų klasifikatoriai su 0,71 F1 reikšme, o ketvirtas geriausiai pasirodęs klasifikavimo algoritmas yra logistinė regresija (F1 reikšmė – 0,69), todėl šie metodai bus bandomi naudojant sudėtinio klasifikatoriaus metodą apjungiant šiuos metodus.

4.4.3.2 Scenarijus, kai tiriama geriausių klasifikatorių kombinacija

Šio scenarijaus metu bus tiriamas geriausiai pasirodžiusių klasifikavimo metodų junginys, kuris susideda iš K artimiausio kaimyno ir atsitiktinių miškų klasifikatorių bei logistinės regresijos klasifikatoriaus. Geriau pasirodęs daugiasluoksnio perceprono klasifikatorius nebuvo suderinamas su kitais klasifikatoriais, todėl buvo naudojamas atsitiktinių miškų klasifikatorius.

Tiriant sudėtinio klasifikatoriaus darbą buvo pastebėta, jog jo tikslumas pagerėja bet labai mažu skirtumu kuris yra 1 arba 2 procentai (žr. 4.22 pav.), lyginant su atskirais klasifikatoriais.

Accuracy Score: 76.3529%

	precision	recall	f1-score
Dos	0.96	0.77	0.86
Normal	0.67	0.97	0.79
Probe	0.79	0.70	0.75
R2L	0.98	0.05	0.09
U2R	0.70	0.28	0.40
accuracy			0.76
macro avg	0.82	0.56	0.58
weighted avg	0.82	0.76	0.72

4.22 pav. Sudėtinio klasifikatoriaus efektyvumo vertinimas, vykdant scenarijų VA

Tame pačiame paveiksle galime pamatyti, jog sudėtinis klasifikatorius labai gerai aptinka DoS atakas (vertinant F1 reikšmę), beveik neaptinka R2L tipo atakų, Tačiau visų šių klasifikatorių blogiausių klasių aptikimo rezultatai yra geresni nei kiekvieno klasifikatoriaus atskirai.

4.4.3.3 Scenarijus, kai tiriamas gilaus mokymosi metodas

Šiame scenarijuje tiriamas „Adam“ gilaus mokymosi metodas, naudojama 20 epochų apsimokymui, naudojama 20 proc. duomenų tikrinimo procesui. Gautus rezultatus galime pamatyti sekančiame paveiksle (žr. 4.23 pav.)

```
Accuracy : 0.8660397444996452
Recall : 0.9424920127795527
Precision : 0.8412742574946095
F1 : 0.8890113928702683

Normal Detection Rate : 0.2349912470394398
Dos Detection Rate : 0.931246726034573
R2L Detection Rate : 0.9243263196751569
U2R Detection Rate : 0.8805970149253731
Probe Detection Rate : 1.0
```

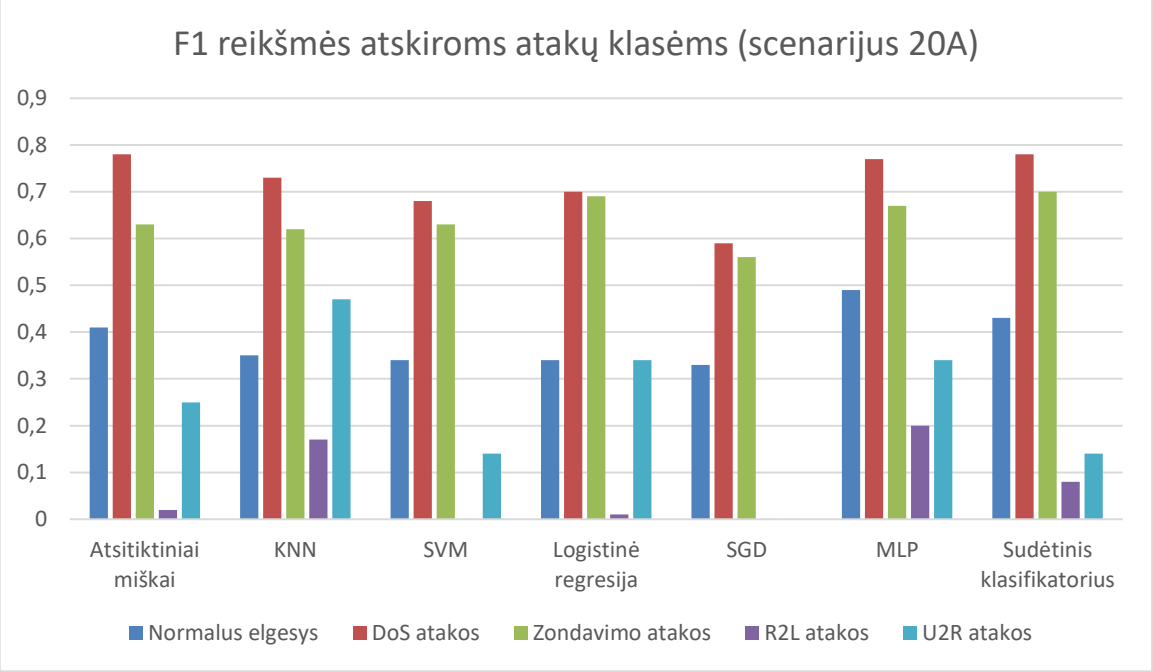
4.23 pav. „Adam“ gilaus mokymosi metodo efektyvumo vertinimas, vykdant scenarijų VA

Iš gautų rezultatų matome, kad visų atakų tipų aptikimo koeficientas yra nuo 88 iki 93 proc., o zondavimo atakos buvo aptiktos 100 proc. ir bendras metodo tikslumas yra 86 proc. (tai yra net 10 procentų daugiau nei sudėtinio klasifikatoriaus gauti rezultatai), tačiau labai prastai buvo aptiktas normalus tinklo elgesys (aptikimo koeficientas tesiekia 23 proc.). Įvertinus šiuos rezultatus galime daryti išvadą, jog gilaus mokymosi metodui buvo pateikta per mažai normalaus elgesio duomenų, kas nulemia mažą šio tipo įrašų klasifikavimo tikslumą.

4.4.4 20A scenarijus (sumažintas duomenų rinkinys ir atrinktos savybės)

4.4.4.1 Pavienių klasifikatorių tyrimas

Šiame scenarijuje buvo naudojamas toks pat duomenų rinkinys kaip ir antrajame tyrimo scenarijuje, tačiau tiriamų savybių kiekis atitinka trečio tyrimo scenarijaus tiriamas duomenų rinkinio savybes.



4.24 pav. Klasifikatorių išvestos F1 reikšmių, įvykdžius scenarijų 20A, diagrama

Išanalizavus klasifikatorių efektyvumo vertinimus (žr. 7.4 lentelė. ir paveikslus nuo 7.1 pav. iki 7.6 pav. bei 4.24 pav.), SVM ir SGD klasifikatoriai nėra tinkami naudojimui nes jie neatpažino 2 iš 5 atakos tipų (R2L ir U2R). Toliau sudarysime sudėtinį klasifikatorių, kurį sudarys 3 geriausiai pasirodę klasifikatoriai: KNN, Atsitiktinių miškų, Logistinės regresijos. Toliau tirsime šį sudarytą klasifikatorių.

4.4.4.2 Scenarijus, kai tiriamas sudėtinis klasifikatorius

Šiame scenarijuje tirtas sudėtinio klasifikatoriaus metodas parametrais atitinka ankstesnį sudėtinio klasifikatoriaus tyrimą (žr. 4.4.3.2 skyrelį).

Accuracy Score: 55.0042%

	precision	recall	f1-score
Dos	0.92	0.68	0.78
Normal	0.28	0.88	0.43
Probe	0.87	0.59	0.70
R2L	0.97	0.04	0.08
U2R	1.00	0.07	0.14
accuracy			0.55
macro avg	0.81	0.45	0.43
weighted avg	0.81	0.55	0.54

4.25 pav. Sudėtinio klasifikatoriaus efektyvumo vertinimas, , vykdant scenarijų 20A

Peržvelgus gautus rezultatus (žr. 4.25 pav.), gauname labai panašias išvadas kaip ir tikrinat ankstesnį sudėtinio klasifikatoriaus tyrimo scenarijų (žr. 4.4.3.2 skyrelį), nes šis klasifikatorius geriausiai pasirodė atpažinant DoS ir zondavimo tipų atakas, tačiau daug prasčiau pasirodė prie normalaus elgesio atpažinimo (F1 reikšmė krito nuo 0,79 iki 0,43) ir U2R tipo atakų (F1 reikšmė sumažėjo nuo 0,40 iki 0,14). Iš šių rezultatų galime padaryti išvadą jog sumažintas duomenų rinkinys nėra tinkamas naudoti apmokant sudėtinio klasifikatoriaus metodą.

4.4.4.3 Scenarijus, kai tiriamas gilaus mokymosi metodas

Šiame scenarijuje naudojami tokie patys parametrai kaip ir ankstesniame „Adam“ metodo tyrimo scenarijuje, tačiau naudojami kitas duomenų rinkinys, kuris yra 20 procentų duomenų paimtų iš pilno rinkinio.

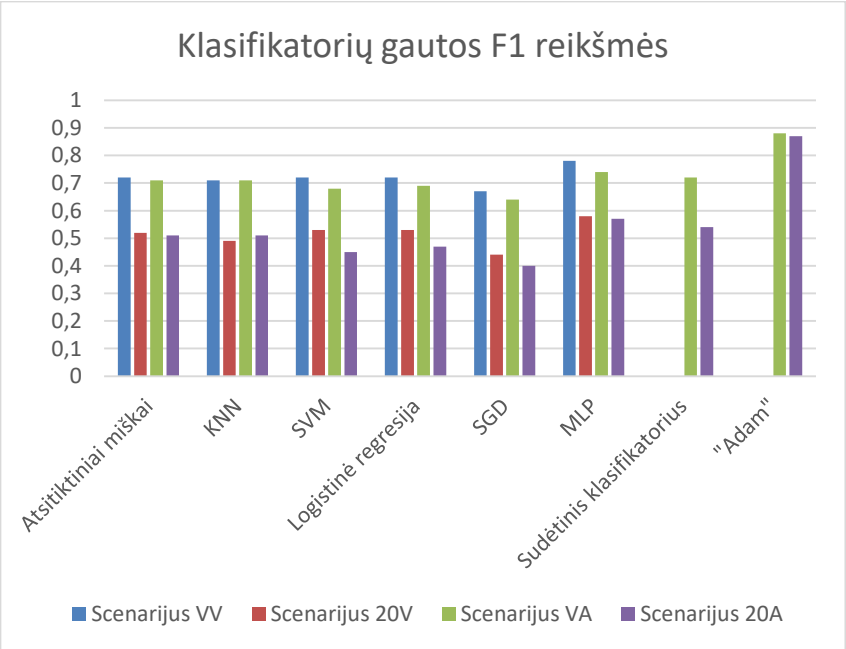
```
Building model for : --- 10.431064128875732 seconds ---
Performance over the testing data set

 Accuracy : 0.790210970464135

 Recall : 0.8592493297587132
 Precision : 0.8814258514914322
 F1 : 0.8701963241436926

Normal Detection Rate : 0.520910780669145
Dos Detection Rate : 0.8555309734513274
R2L Detection Rate : 0.7379106681432263
U2R Detection Rate : 0.9701492537313433
Probe Detection Rate : 1.0
```

4.26 pav. „Adam“ gilaus mokymosi metodo efektyvumo vertinimas, vykdant scenarijų 20A

Iš rezultatų paveikslo (žr. 4.26 pav.) galime pamatyti, jog šiuo duomenų rinkiniu apmokius metodą, bendras jo tikslumas siekia 79 procentus. Giliau pažvelgę galime pamatyti jos U2R ir zondavimo tipo atakas, metodas aptiko beveik tobulai (atitinkamai 97 ir 100 procentai tikslumu). Tačiau kaip ir naudojant pilną duomenų rinkinį, normalaus tipo įrašus aptinka žemu tikslumu (tik 52 procentai). Apžvelgus gautus tyrimo scenarijaus rezultatus galime teigti, kad kaip ir praeito scenarijaus atveju yra per mažai normalaus tipo duomenų įrašų, kas lema sunkiai atpažintus normalaus elgesio įrašus.

4.5 Tyrimo išvados

Eksperimento metu gauti ir palyginti rezultatai su keturiais skirtingais duomenų rinkinio panaudojimo scenarijais (žr. 4.27 pav.).



4.27 pav. Apibendrinti tyrimo rezultatai

Išanalizavus gautus tyrimo rezultatus galime pamatyti, jog efektyviausi atakų klasifikavimo yra KNN, MLP, atsitiktinių miškų ir logistinės regresijos metodai. Tačiau daugiasluoksnio perceprono metodas nesusiderina su kitais metodais sudėtiniame klasifikatoriuje, todėl vietoje jo buvo naudojamas atsitiktinių miškų klasifikatorius.

Eksperimentais nustatyta, kad analizuoti klasifikatoriai geriausiai aptinka DoS tipo atakas (75-80%), zondavimo tipo atakas aptinka prasčiau (60-70%) ir duoda labai skirtingus rezultatus R2L bei U2R tipo atakoms (0-40%).

Sudarius sudėtinį klasifikatorių iš anksčiau minėtų 3 metodų buvo gauti geresni rezultatai, nei kiekvieno metodo atskirai, kurie yra 70-74 proc., sudėtinio klasifikatoriaus tikslumas yra 76,35 proc.

Sulyginus sudėtinio klasifikatoriaus su gilaus mokymosi metodo gautais duomenimis buvo pastebėta, kad net ir apjungus geriausiai pasirodžiusius klasifikatorius į vieną, jo rezultatai neprilygo gilaus mokymosi metodui, nes atakų aptikimo tikslumas paliko nuo 76,35 proc. iki 86,6 proc. Tačiau gilaus mokymosi metodas su tais pačiais duomenimis labai sunkiai aptiko normalaus elgesio įrašus („Adam“ optimizatoriaus metodas aptinka tik 24 proc. normalaus elgesio įrašų, kai sudėtinis klasifikatorius net 97 proc.).

Sumažinus duomenų rinkinio savybių skaičių, rezultatai suprastėjo labai mažu kiekiu (1-2%). Tai reiškia, kad pašaliname nereikalingas savybes,

Sumažinus duomenų rinkinio dydį 5 kartus gavome 20% tikslumo sumažėjimą.

Iš gilaus mokymosi metodo rezultatų galime teigti, jog šis duomenų rinkinys netinka gilaus mokymosi metodų taikymui, dėl nesubalansuoto atakų tipų įrašų skaičiaus.

Duomenų rinkinio trūkumus galime pastebėti ir tiriant jį su klasifikatoriais, kurie buvo įvertinti aptinkant R2L tipo atakas laba žema F1 reikšme (pvz. 0; 0,01; 0,09), todėl šis duomenų rinkinys nėra tinkamas šio tipo atakoms atpažinti.

5 DARBO IŠVADOS

1. Atlikus šaltinių analizę, buvo pastebėta, jog tinklo anomalijų aptikimas yra viena iš pagrindinių tinklo saugos priemonių, tačiau naudojamų metodų efektyvumas labai varijuoja priklausomai nuo tinklo srauto ypatybių ir atakų savybių. Todėl reikalingas apsimokymas naudojant atitinkančius tinklo situaciją duomenis. Darbe iškeltas uždavinys ištirti apsimokymo metodų efektyvumą naudojant Cisco Netflow srauto įrašus.

2. Suprojektuota apsimokymo metodų efektyvumo tyrimo sistema, kurios išskirtinėmis savybėmis yra:

– galimybė manipuliuoti apmokymo rinkinio savybėmis;
– galimybė tirti tiek pavienius tiek sudėtinius klasifikatorius.

3. Apart klasifikatorių efektyvumo tyrimo, eksperimentų metu siekiama nustatyti kaip įtakoja rezultatus:

– apsimokymo rinkinio dydis;
– apsiribojimas tik tomis rinkinio savybėmis, kurios teoriškai labiausiai būdingos pasirinktų atakų tipui ir gali būti išgaunamos iš „Cisco Netflow“ rinkinio.

4. Realizuoti prototipai, kurie naudoja NSL-KDD sintetinio duomenų rinkinio duomenis. Įgyvendinti 6 klasifikatoriai (atsitiktinių miškų, KNN, SVM, logistinės regresijos, SGD ir MLP klasifikatoriai), kurių įverčiais remiantis bus sudarytas galutinis sudėtinio klasifikatoriaus modelis. Pagal turimą duomenų rinkinį tyrimas orientuotas keturių tipų atakų (DoS, zondavimo, R2L ir U2R) atpažinimui.

5. Atlikto tyrimo rezultatai parodė, kad apsimokymui sumažinus imtį penkis kartus rezultatai suprastėjo apie 20%. Siekiant didesnio tikslumo, mažinti duomenų rinkinio nereikėtų.

6. Ištyrus klasifikatorių veikimą su sumažintu „Cisco Netflow“ rinkinio savybių kiekiu gauti rezultatai iš esmės nepablogėjo, todėl galime teigti jog buvo pašalintos tik neesminės savybės.

7. Ištyrus atskirus klasifikatorius buvo pastebėta jos efektyviausiai dirba KNN, MLP, atsitiktinių miškų ir logistinės regresijos metodai (atitinkamai F1 reikšmės yra 0,71;0,74;0,71;0,69).

8. Tyrimas parodė, kad sudėtinis klasifikatorius pagerino efektyvumą DoS ir zondavimo atakų tipams.

9. Kontrolinis gilaus mokymosi „Adam“ metodas parodė prieštaringus rezultatus. Atakų įrašus jis atpažįsta tiksliau (86% prieš 76%), tačiau labai blogai identifikuoja normalų srautą (24%). Tokio elgesio priežasčių nustatyti nepavyko.

10. Galima teigti jog NSL-KDD duomenų rinkinyje nėra subalansuoto kiekio duomenų, kad pakankamu tikslumu būtų atpažįstamos U2R ir R2L tipo atakos. „Cisco Netflow“ įrašai netinkami šių atakų atpažinimui, nes neturi reikalingų požymių.

6 LITERATŪRA

1. HAMAMOTO, A.H. ir kt. Network Anomaly Detection System using Genetic Algorithm and Fuzzy Logic. In *Expert Systems with Applications* [interaktyvus]. 2018. Vol. 92, p. 390–402. [žiūrėta 2020-01-12]. . Prieiga per internetą: <http://www.sciencedirect.com/science/article/pii/S095741741730619X>.

2. YASAMI, Y. - MOZAFFARI, S.P. A novel unsupervised classification approach for network anomaly detection by k-Means clustering and ID3 decision tree learning methods. In *The Journal of Supercomputing* [interaktyvus]. 2010. Vol. 53, no. 1, p. 231–245. [žiūrėta 2019-11-25]. . Prieiga per internetą: <http://link.springer.com/10.1007/s11227-009-0338-x>.

3. BARFORD, P. - PLONKA, D. Characteristics of Network Traffic Flow Anomalies. In . p. 5. .

4. Allot Anomaly Detection | AllotWorks.com. In [interaktyvus]. [žiūrėta 2019-11-25]. Prieiga per internetą: <http://www.allotworks.com/Anomaly-Detection.asp>.

5. MOUSTAFA, N. ir kt. A holistic review of Network Anomaly Detection Systems: A comprehensive survey. In *Journal of Network and Computer Applications* . 2018. Vol. 128. .

6. LATAH, M. - TOKER, L. An Efficient Flow-based Multi-level Hybrid Intrusion Detection System for Software-Defined Networks. In . p. 14. .

7. KOTPALLIWAR, M.V. - WAJGI, R. Classification of Attacks Using Support Vector Machine (SVM) on KDDCUP'99 IDS Database. In *2015 Fifth International Conference on Communication Systems and Network Technologies* . 2015. p. 987–990. .

8. AMIRGHOLIPOUR KASMANI, S. - POUREBRAHIMI, A. Intrusion Detection Based on Joint of K-Means and KNN. In *Journal of Convergence Information Technology(JCIT)* . 2015. Vol. 10, p. 42–51. .

9. GAO, N. ir kt. An Intrusion Detection Model Based on Deep Belief Networks. In *2014 Second International Conference on Advanced Cloud and Big Data* . 2014. p. 247–252. .

10. INGRE, B. ir kt. Decision Tree Based Intrusion Detection System for NSL-KDD Dataset. In . 2017. .

11. BEGGEL, L. ir kt. Robust Anomaly Detection in Images using Adversarial Autoencoders. In . p. 16. .

12. Network-Wide Traffic Anomaly Detection and Localization Based on Robust Multivariate Probabilistic Calibration Model. In [interaktyvus]. [žiūrėta 2019-12-02]. Prieiga per internetą: <https://www.hindawi.com/journals/mpe/2015/923792/>.

13. KURNIABUDI, K. ir kt. Network anomaly detection research: a survey. In *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)* [interaktyvus]. 2019. Vol. 7, no. 1, p. 37–50. [žiūrėta 2019-11-25]. . Prieiga per internetą: <http://section.iaesonline.com/index.php/IJEEI/article/view/773>.

14. Cisco Stealthwatch. In *Cisco* [interaktyvus]. [žiūrėta 2020-01-19]. Prieiga per internetą: <https://www.cisco.com/c/en/us/products/security/stealthwatch/index.html>.

15. IBM QRadar Security Intelligence. In [interaktyvus]. 2020. [žiūrėta 2020-01-19]. Prieiga per internetą: <https://www.ibm.com/security/security-intelligence/qradar>.

16. McAfee Network Threat Behavior Analysis 9.1.x Product Guide. In . p. 141. .

17. [Interaktyvus]. 2018. [žiūrėta 2020-01-19]. Prieiga per internetą: <https://www.lastline.com/solutions/network-defender/>.

18. Ourmon network monitoring and anomaly detection system. In [interaktyvus]. [žiūrėta 2020-01-20]. Prieiga per internetą: <http://ourmon.sourceforge.net/>.

19. *Tripwire/tripwire-open-source* [interaktyvus]. . [s.l.]: Tripwire Inc., 2020. .

20. Security Onion. In [interaktyvus]. [žiūrėta 2020-01-20]. Prieiga per internetą: <https://securityonion.net/>.

21. Netflow - What is it, a Definition & How to Collect & Analyze Flow Data (sFlow, Ipfix, jFlow, etc). In *Software Portal* [interaktyvus]. 2019. [žiūrėta 2021-01-16]. Prieiga per internetą: <https://softwareportal.com/netflow/>.

22. Network as a Security Sensor White Paper. In *Cisco* [interaktyvus]. [žiūrėta 2021-01-16]. Prieiga per internetą: <https://www.cisco.com/c/en/us/solutions/collateral/enterprise-networks/enterprise-network-security/white-paper-c11-736595.html>.

23. DAMASEVICIUS, R. ir kt. LITNET-2020: An Annotated Real-World Network Flow Dataset for Network Intrusion Detection. In *Electronics* . 2020. Vol. 9, p. 800. .

24. Published on IST 554 (https://online.edu/ist554) Topic 7: Intrusion Detection Today's highly connected computing environment offers unlimited... | Course Hero. In [interaktyvus]. [žiūrėta 2021-01-20]. Prieiga per internetą: <https://www.coursehero.com/tutors-problems/Computer-Science/8448801-A-Phf-attack-is-a-remote-to-local-R2L-attack-against-the-Web-Server/>.

25. DEY, D. ir kt. Warezmaster and Warezclient: An implementation of FTP based R2L attacks. In *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* . 2017. p. 1–6. .

26. [Interaktyvus]. 2020. [žiūrėta 2021-01-20]. Prieiga per internetą: <https://en.wikipedia.org/w/index.php?title=Rootkit&oldid=995296552>.

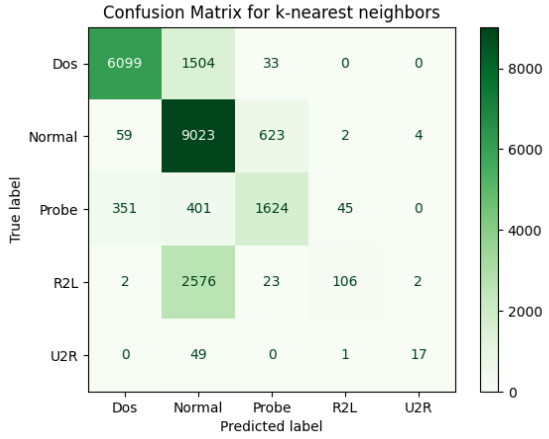
27. NSL-KDD | Datasets | Research | Canadian Institute for Cybersecurity | UNB. In [interaktyvus]. [žiūrėta 2020-04-19]. Prieiga per internetą: <https://www.unb.ca/cic/datasets/nsl.html>.

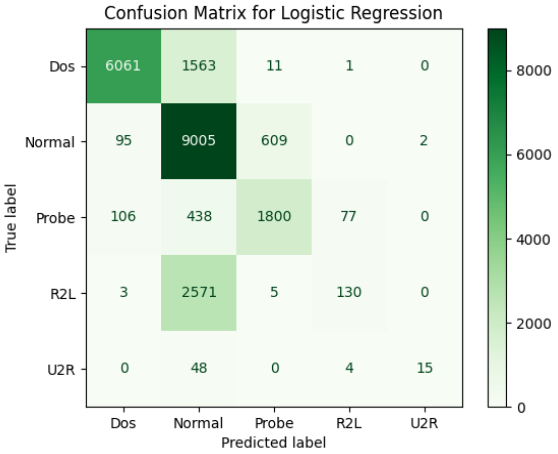
28. SAPORITO, G. A Deeper Dive into the NSL-KDD Data Set. In *Medium* [interaktyvus]. 2019. [žiūrėta 2020-09-22]. Prieiga per internetą: <https://towardsdatascience.com/a-deeper-dive-into-the-nsl-kdd-data-set-15c753364657>.

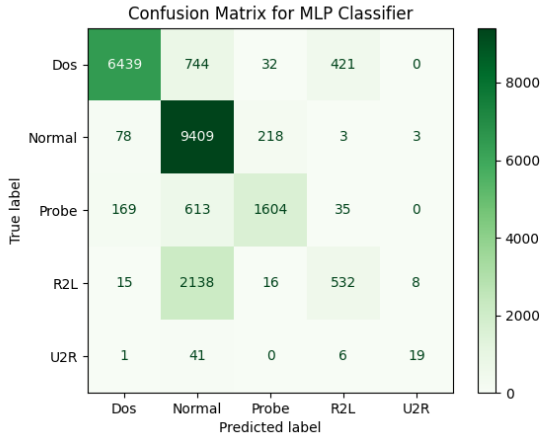
29. NOURELDIEN, N.A. - YOUSIF, I.M. Accuracy of Machine Learning Algorithms in Detecting DoS Attacks Types. In *Science and Technology* . 2016. Vol. 6, no. 4, p. 89–92. [žiūrėta 2020-06-16]. . .

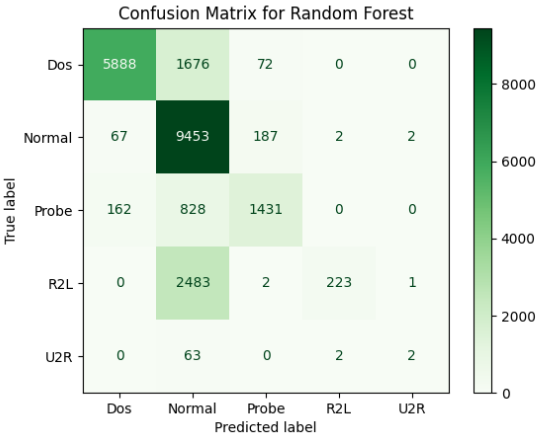
7 PRIEDAI

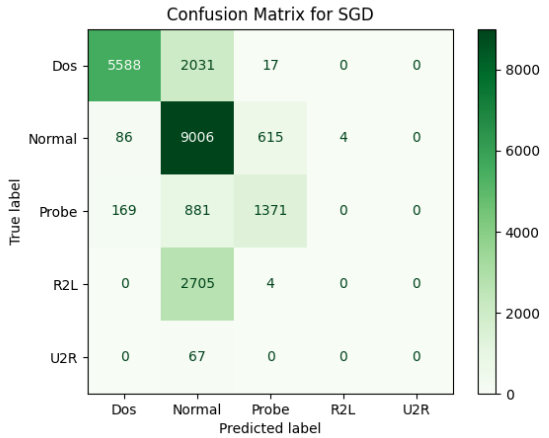
7.1 priedas. Scenarijaus VV išvestys.

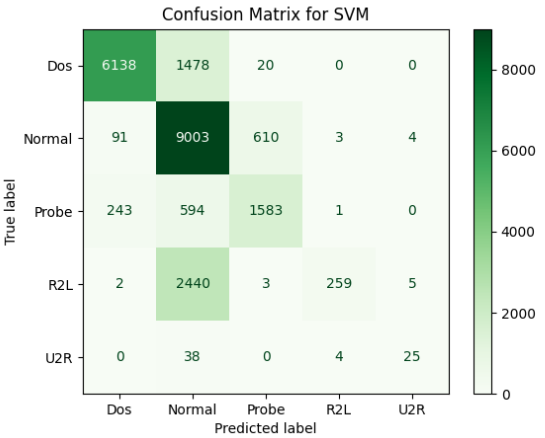








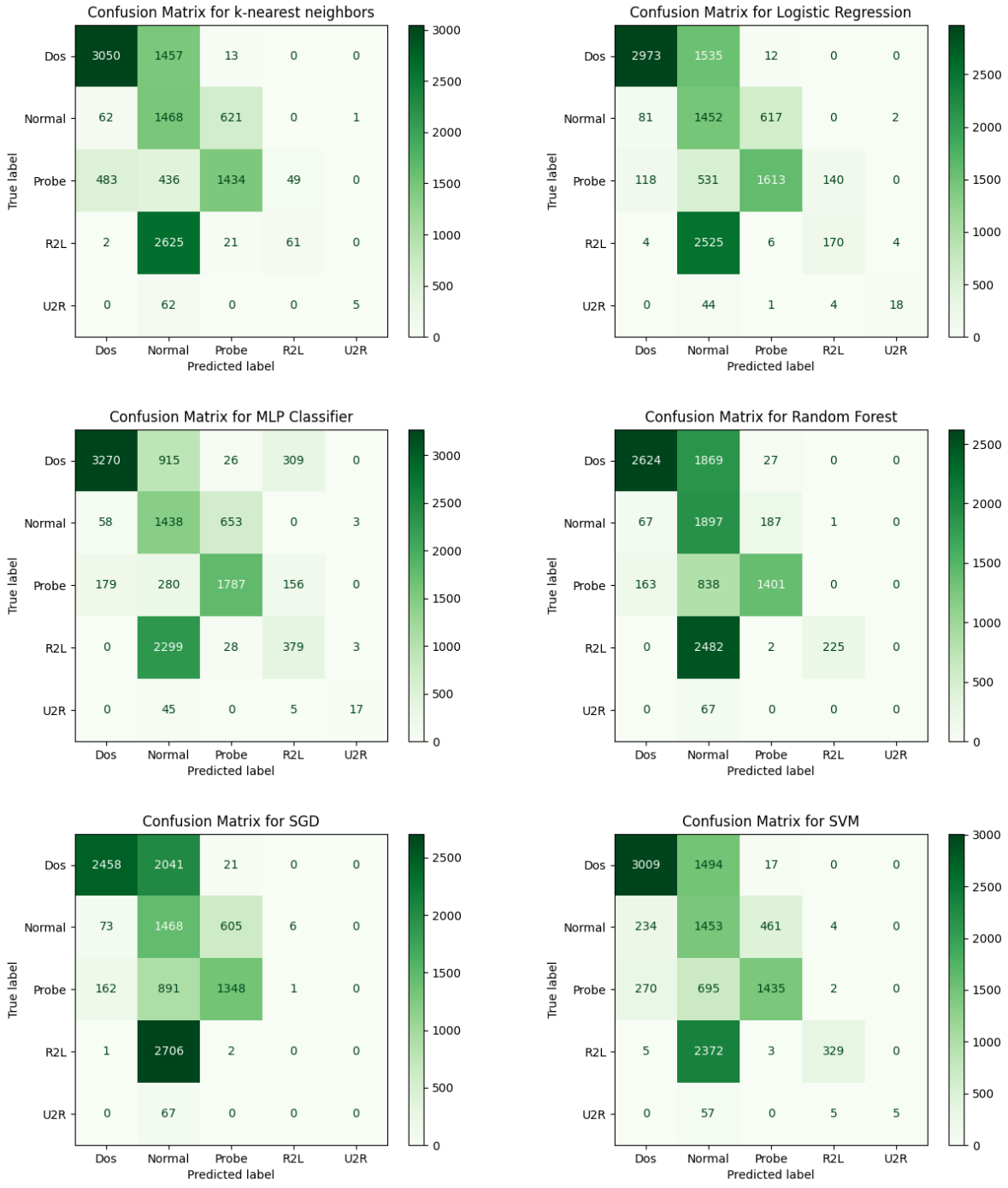




7.1 lentelė. Klasifikatorių efektyvumo vertinimas scenarijuje VV

Metodo pavadinimas	**Tikslumas**	**Aptikimo koeficientas**	**Preciziškumas**	**F1 reikšmė**
Atsitiktiniai miškai	0.7539	0.75	0.82	0.72
KNN	0.7482	0.75	0.76	0.71
SVM	0.7544	0.75	0.80	0.72
Logistinė regresija	0.7545	0.75	0.77	0.72
SGD	0.7081	0.71	0.66	0.67
MLP	0.7985	0.80	0.80	0.78

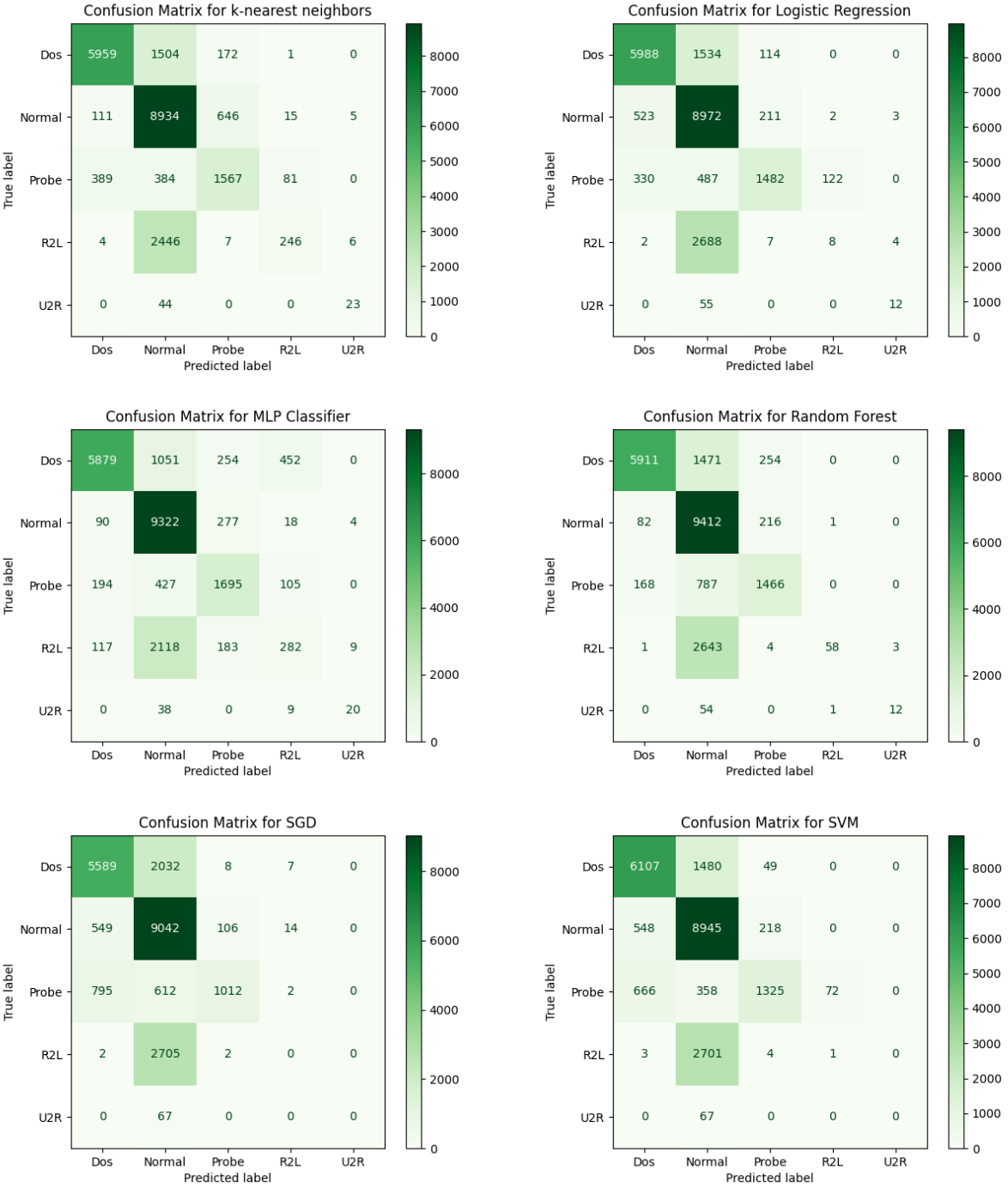
7.2 priedas. Scenarijaus 20V išvestys.

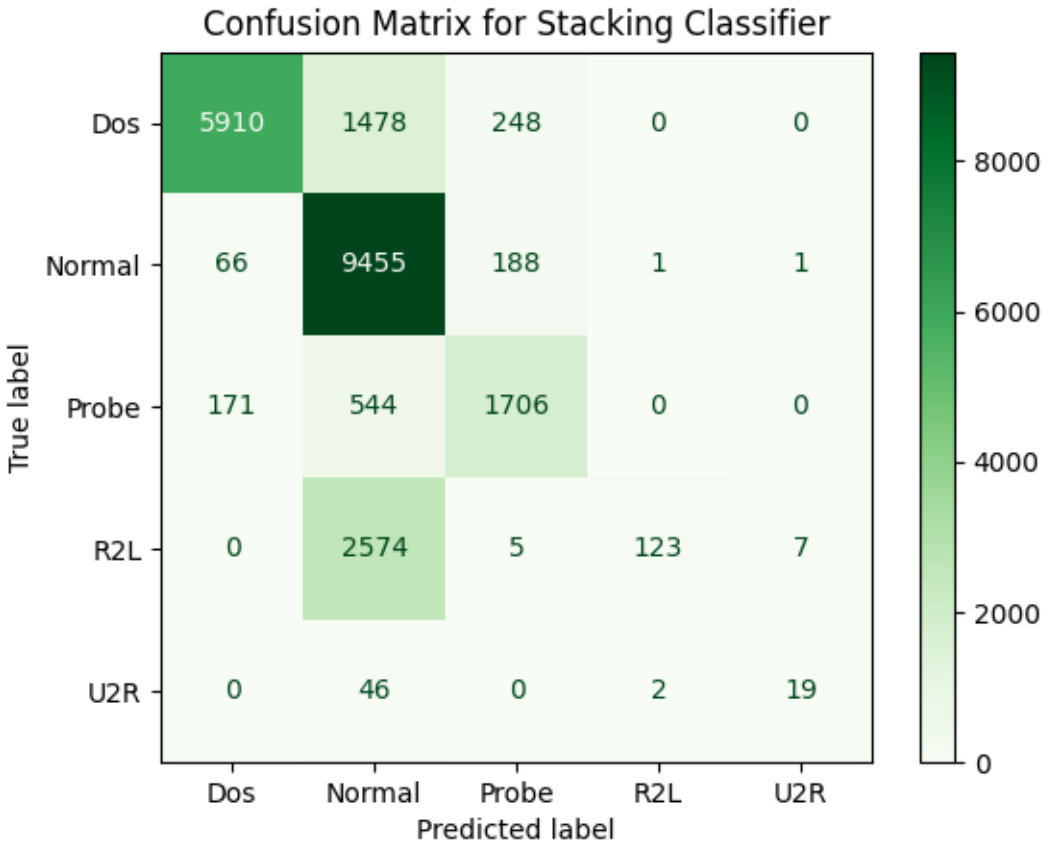


7.2 lentelė. Klasifikatorių efektyvumo vertinimas scenarijuje 20V

Metodo pavadinimas	**Tikslumas**	**Aptikimo koeficientas**	**Preciziškumas**	**F1 reikšmė**
Atsitiktiniai miškai	0.5187	0.52	0.80	0.52
KNN	0.5078	0.51	0.64	0.49
SVM	0.5258	0.53	0.75	0.53
Logistinė regresija	0.5254	0.53	0.67	0.53
SGD	0.4450	0.45	0.52	0.44
MLP	0.5815	0.58	0.66	0.58

7.3 priedas. Scenarijaus VA išvestys.

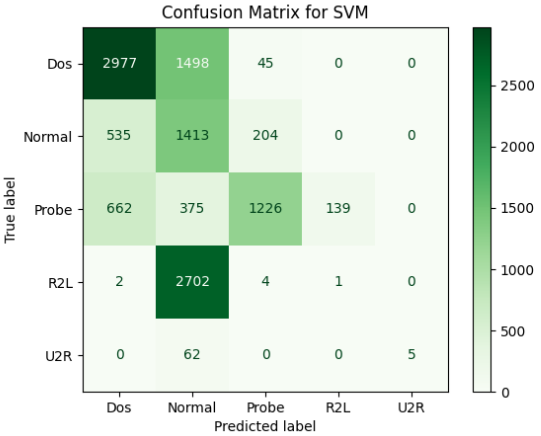


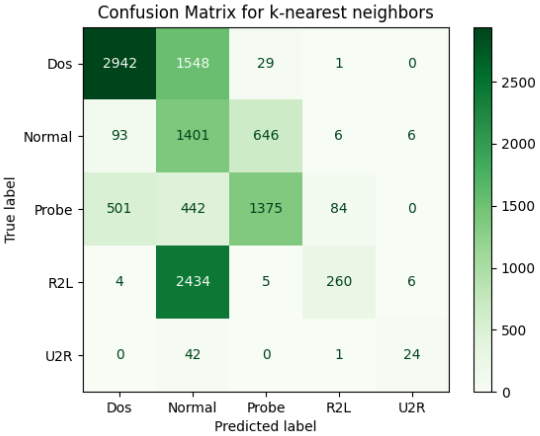


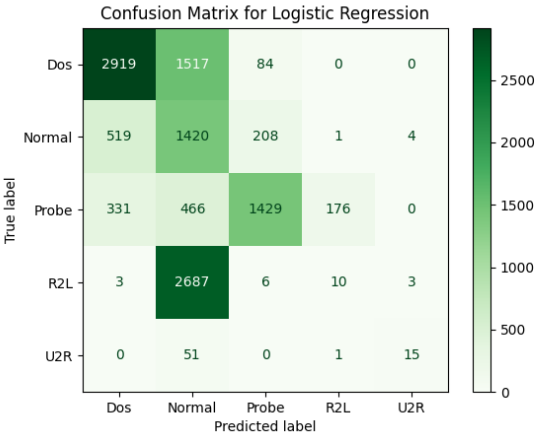
7.3 lentelė. Klasifikatorių efektyvumo vertinimas scenarijuje VA

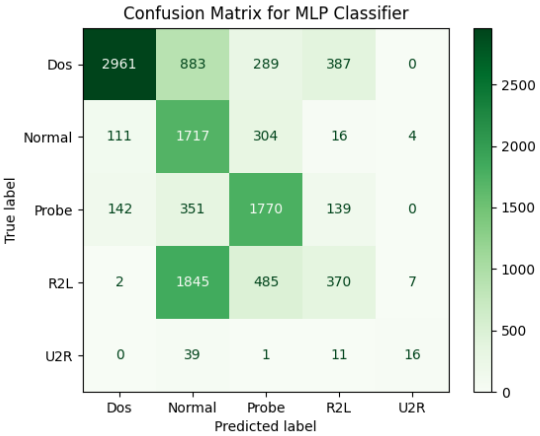
Metodo pavadinimas	**Tikslumas**	**Aptikimo koeficientas**	**Preciziškumas**	**F1 reikšmė**
Atsitiktiniai miškai	0.7478	0.75	0.81	0.71
KNN	0.7420	0.74	0.76	0.71
SVM	0.7264	0.73	0.66	0.68
Logistinė regresija	0.7302	0.73	0.67	0.69
SGD	0.6938	0.69	0.64	0.64
MLP	0.7628	0.76	0.74	0.74

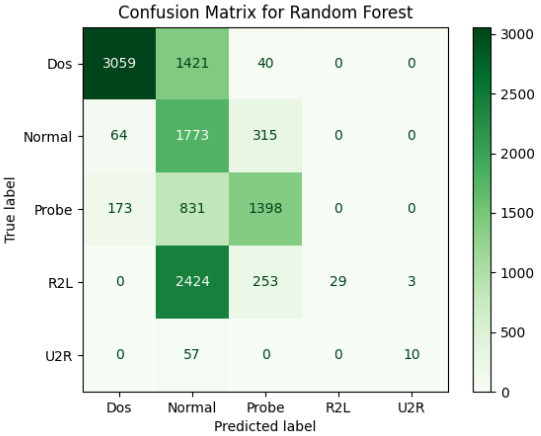
7.4 priedas. Scenarijaus 20A išvestys.

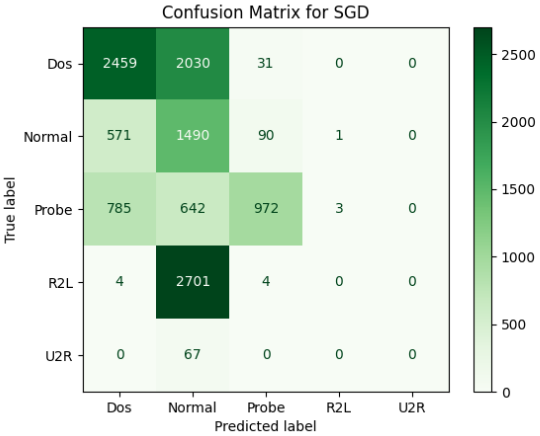












	precision	recall	f1-score
Dos	0.96	0.77	0.86
Normal	0.66	0.97	0.78
Probe	0.76	0.61	0.67
R2L	0.97	0.02	0.04
U2R	0.80	0.18	0.29

7.1 pav. Atsitiktinių miškų metodo rezultatai atskiroms atakų klasėms, kai vykdome scenarijų 20A

	precision	recall	f1-score
Dos	0.92	0.78	0.85
Normal	0.67	0.92	0.78
Probe	0.66	0.65	0.65
R2L	0.72	0.09	0.16
U2R	0.68	0.34	0.46

7.2 pav. KNN metodo rezultatai atskiroms atakų klasėms, kai vykdome scenarijų 20A

	precision	recall	f1-score
Dos	0.83	0.80	0.82
Normal	0.66	0.92	0.77
Probe	0.83	0.55	0.66
R2L	0.01	0.00	0.00
U2R	0.00	0.00	0.00

7.3 pav. SVM metodo rezultatai atskiroms atakų klasėms, kai vykdome scenarijų 20A

	precision	recall	f1-score
Dos	0.88	0.78	0.83
Normal	0.65	0.92	0.77
Probe	0.82	0.61	0.70
R2L	0.06	0.00	0.01
U2R	0.63	0.18	0.28

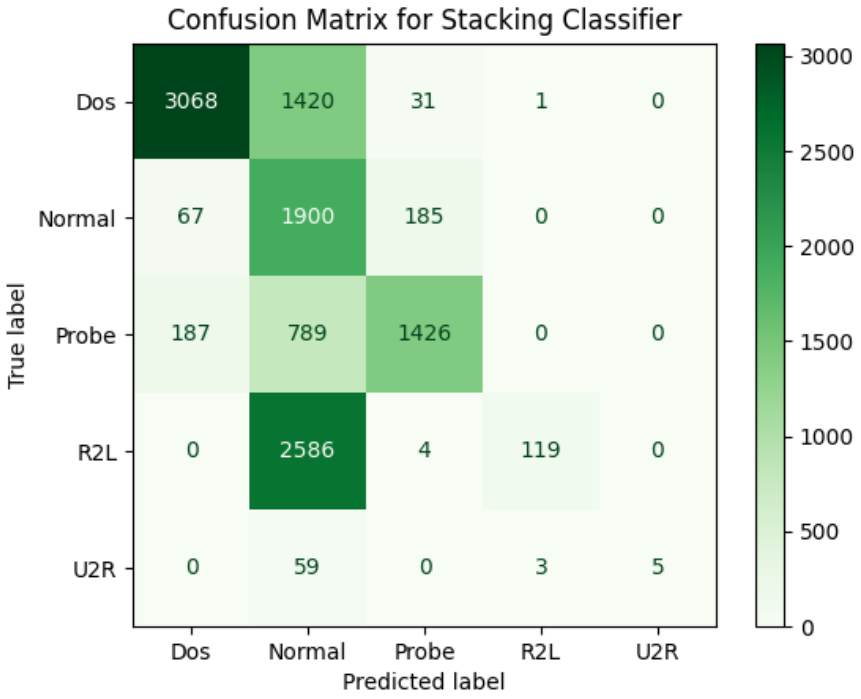
7.4 pav. Logistinės regresijos metodo rezultatai atskiroms atakų klasėms, kai vykdome scenarijų 20A

	precision	recall	f1-score
Dos	0.81	0.73	0.77
Normal	0.63	0.93	0.75
Probe	0.90	0.42	0.57
R2L	0.00	0.00	0.00
U2R	0.00	0.00	0.00

7.5 pav. SGD metodo rezultatai atskiroms atakų klasėms, kai vykdome scenarijų 20A

	precision	recall	f1-score
Dos	0.94	0.77	0.84
Normal	0.72	0.96	0.82
Probe	0.70	0.70	0.70
R2L	0.33	0.10	0.16
U2R	0.61	0.30	0.40

7.6 pav. MLP metodo rezultatai atskiroms atakų klasėms, kai vykdome scenarijų 20A



7.4 lentelė. Klasifikatorių efektyvumo vertinimas, vykdant scenarijų 20A

Metodo pavadinimas	**Tikslumas**	**Aptikimo koeficientas**	**Preciziškumas**	**F1 reikšmė**
Atsitiktiniai miškai	0.5290	0.53	0.78	0.51
KNN	0.5064	0.51	0.67	0.51
SVM	0.4744	0.47	0.49	0.45
Logistinė regresija	0.4888	0.52	0.49	0.47
SGD	0.4152	0.42	0.46	0.40
MLP	0.5767	0.58	0.64	0.57