Vilniaus universitetas
Duomenų mokslo ir skaitmeninių technologijų institutas
L I E T U V A

INFORMATIKA (N009)

MAŠININIO MOKYMO METODŲ EFEKTYVUMO TYRIMAS SPRENDŽIANT ANKSTYVOJO KIBERNETINIŲ INCIDENTŲ APTIKIMO UŽDAVINĮ

Viktoras Bulavas

2020 m. spalis

Mokslinė ataskaita DMSTI-DS-N009-20-07

VU Duomenų mokslo ir skaitmeninių technologijų institutas, Akademijos g. 4,
Vilnius LT-08412
www.mii.lt

Santrauka

Šioje ataskaitoje pateikiami atliktų teorinių ir empirinių tyrimų rezultatai, kurių tikslas – parinkti tinkamus mašininio mokymo algoritmus ir suderinti jų hiperparametrus sprendžiant kibernetinių incidentų ankstyvojo aptikimo uždavinį.

Pirmame tyrimo etape atlikta duomenų šaltinio parinkimo analizė ir parengimas mašininio mokymo algoritmams. Pasirinktas kibernetinio saugumo duomenų šaltinio pilnumo kriterijus atitinkantis atviros prieigos kibernetinės saugos duomenų rinkinys CIC-IDS-2017 (Sharafaldin, Lashkari and Ghorbani, 2018), kuriame yra dokumentuoti skirtingi kibernetinio įsibrovimo incidentai.

Antrame etape duomenys buvo nagrinėti siekiant išsiaiškinti svarbiausius požymius, įnešančius didžiausią indėlį į mašininio mokymo tikslumą. Papildant ankstesnių metų tyrimus, kuriuose požymiai nagrinėti pagrindinių komponenčių (PCM) ir Šapo verčių nustatymo metodais, išbandyti dar trys skirtingi būdai parenkant savybes: koreliacijų analizė, reikšmingiausių savybių atrinkimas KBest metodu ir rekursinis reikšmingiausių savybių atrinkimas (RFE).

Trečiame tyrimo etape, papildant duomenų šaltinių agregatų sąrašą, pasiūlytas papildomas statistinis agregatas, fraktalinė dimensjia. Greta klasikinių statistinių metodų kibernetinės atakos atpažinimui buvo išbandytas T. Higuchi fraktalinės dimensijos skaičiavimo Boxcount metodas (Higuchi, 1988). Šio metodo taikymą tinklo srauto vertinimui pirmieji aprašė Xia, Lu ir Li (Xia, Lu and Li, 2012). Šiame tyrime remiantis tinklo srauto fraktalinės dimensijos kaitos laike skaičiavimais buvo gauti gerai faktinius duomenis atitinkantys rezultatai, tyrimo rezultatai pristatyti konferencijoje.

Ketvirtame etape Python aplinkoje su CIC-IDS-2017 duomenimis realizuoti lyginamojo duomenų šaltinio autorių publikuoto tyrimo mašinų mokymo algoritmai, ir palygintas gautas efektyvumas. Siekiant šio tikslo, panaudojant tinklelio paieškos (‚Grid Search‘) ir kryžminio patikrinimo ('cross-validation’) metodus atrinkti skirtingų algoritmų mokymo parametrai. Suskaičiuoti algoritmo tikslumo (accuracy), klasių prognozės tikslumo (precision), jautrumo (recall), harmoninio tikslumo F1, Hamming Loss ir Jaccart Score, Balanced Accuracy Score rodikliai. Publikacija rengiama, todėl pateikiami tarpinių tyrimų rezultatai gali neatitikti rengiamų publikacijai.

Reikšminiai žodžiai: kibernetinė sauga; incidentų atpažinimo metodai; mašininis mokymas; dedamųjų atranka; hiperparametrai.

Turinys

1 Uždavinio aktualumo aptarimas

Ankstyvas įsibrovimo į tinklą, informacines sistemas ar darbo vietas atpažinimas, bei susijęs kenksmingos programinės įrangos operacijų ar nesankcionuotos individų veiklos aptikimas yra aktualūs kibernetinio saugumo srities uždaviniai. Įsibrovimų į tinklus, informacines sistemas ar darbo stotis nustatymas, taip pat kenkėjiškų programų ir asmenų neteisėtos veiklos aptikimas tapo globaliu iššūkiu. Kompiuterinio raštingumo augimas ir lengvėjanti prieiga prie kompiuterinių technologijų sudaro prielaidas nusikaltimams, kurių nesulaiko valstybių sienos. Pradėjus kibernetinių išpuolių tyrimus, prireikė tam skirtų, veikiančių visą parą, specialiųjų pajėgų. Atsiradus besimokančioms kibernetinėms puolimo sistemoms, įsilaužimų prevencijos sistemos retai besustabdo tokius užpuolikus. Patiems piliečiams trūksta specializuotų žinių, brangios programinės įrangos ir laiko, jie patys retai apsigina nuo kibernetinių atakų. Reaguodama į šią problemą, NATO 2016 m. Varšuvos aukščiausiojo lygio susitikime paskelbė kibernetinę erdvę operacijų sritimi, tokia kaip ir oras, sausuma ir jūra. Visa paminėta patvirtina labai reikalingą tyrimų sritį, kuria siekiama padidinti grėsmių nustatymo tikslumą ir susijusio reagavimo greitį.

Šiuos iššūkius gali padėti įveikti mašininiu mokymu paremtų įsilaužimo aptikimo algoritmų pritaikymas, kuris yra šios studijos tyrimo objektas.

2 Tyrimuose naudotų duomenų aptarimas

2.1 Kibernetinių įvykių duomenų formatai

Nors dominuoja tinklo įrangos gamintojų rinkos lyderių formatai, kurie nuolat keičiasi su naujais įrangos modeliais ir technologijomis, tokiomis, kaip duomenų šifravimas, kyla poreikis generuoti naujus mokslui tinkamus duomenų šaltinius. Tiekėjai teikia panašią, bet nevienodą tinklo srautų stebėjimo technologiją, nes technologijos yra patentų objektas. Todėl algoritmų mokymas turi būti pritaikytas konkrečiam srautų įrašų tipui arba perdirbamas specialiomis priemonėmis. Mokslinėse duomenų bazėse kalbama apie Netflows, kam pritaikyti dabartiniai atviro kodo sprendimai, tačiau jau plinta naujesnis tinklo srauto formatas IPFIX. Todėl įsibrovimo aptikimo sistemos turės būti pritaikytos arba pakeistos.

Šiame tyrime naudojami 3 rūšių tinklo duomenų formatai:

a) CIC-FlowMeter[1] priemone agreguoti, sodrinti ir žymėti tinklo veiklos įrašai, pateikiami 1 lentelėje.
b) PCAP[2] - pilnos srauto įrenginių nuskaitymo būdu sukauptos tinklo veiklos srauto kopijos išranka, lentelės 2 ir 3;
c) NetFlows[3] – agreguoti tinklo veiklos srautai;

PCAP duomenys panaudoti iš CIC-IDS-2018 duomenų bazės ir autoriaus atliktų Vilniaus universiteto tinklo stebėjimų, NetFlows duomenys autoriaus generuoti iš PCAP, pateiktų su CIC-IDS-2018 duomenų baze.

2.2 Duomenų šaltiniai

Mokslo bendruomenė, tyrimų atkartojimo tikslais, naudoja apjungtus su sisteminiais ir nuasmenintus kibernetinės saugos duomenų rinkinius. Paminėtini Linkolno universiteto surinkti DARPA 1998 (Laboratory, 1998) ir vėlesnių metų rinkiniai, KDD'99 (KDD, 1999), Kalifornijos universiteto, interneto srauto archyvas CAIDA (The Cooperative Association for Internet Data Analysis, 2010), taip pat Lawrence Berkeley National Laboratory, JAV archyvas LBNL (Lawrence Berkeley National Laboratory, 2010), „Shmoo Group“, JAV Defcon archyvas (The Shmoo Group, 2011), Kanados kibernetinio saugumo instituto ir Naujojo Brunšviko universiteto archyvai ISCX IDS 2012 ir CIC IDS 2017, CSE-CIC-IDS2018 (Sharafaldin et al., 2018), ir kiti.

(Gharib et al., 2016) tyrime pateiktas išsamus kriterijų, apibūdinančių kibernetinio saugumo mokslinių tyrimų duomenų šaltinių poreikius, aprašymas. Autoriai įvardina, kad moksliniams tyrimams kibernetinio saugumo srityje skirtas duomenų šaltinis turėtų tenkinti žemiau įvardintus vienuolika kriterijų. Nė vienas iš ankstesnių IDS duomenų rinkinių neatitinka visų šių kriterijų, todėl tyrimui pasirinkti būtent CIC-IDS-2017 ir CIC-IDS-2018 duomenų rinkiniai. Toliau aptariami šie vienuolika kriterijų.

1. Pilna tinklo konfigūracija: Norint tirti tikrovišką atakų eigą, būtina išbandyti tikrovišką tinklo konfigūraciją. Tuo būdu sudaromas realistinės organizacijos įrangos konfigūracijos duomenų įrašų kaupimo poligonas, kuriame pratybų metu atliekamos kibernetinių atakų simuliacijos.

[1] CIC-FlowMeter - Copyright (c) 2016 Canadian Institute for Cybersecurity (CIC)
[2] PCAP – sutrumpinimas nuo angl. Packet CAPture
[3] NetFlows – sutrumpinimas nuo ang. Network Flows

2. Reprezentatyvi infrastruktūra: kaupiami pilni paketų iš šaltinio iki paskirties taško, kuris gali būti darbo vietos kompiuteris, maršrutizatorius, kitas įrenginys, specializuotos tarnybos įrenginys, srautai.

3. Paženklintas duomenų rinkinys: svarbu turėti patikimas srautų etiketes, kurios būtų naudingesnės ir suprantamesnės vartotojams. Srautai turi būti sužymėti bent jau į gerybinius ar piktybinius.

4. Visiška sąveika: norint teisingai interpretuoti duomenis, reikia visos tinklo sąveikos proceso duomenų.

5. Įrašų pilnumas: būtina turėti įrašus, net jei srautas neuždarytas. Kai kuriuose duomenų rinkiniuose neuždarytų srautų informacija pašalinama. Tai svarbu skaičiuojant atpažinimo tikslumo rodiklius.

6. Įvairūs protokolai: reikia simuliuoti kiek galima tikroviškesnio normalaus naudotojų elgesio įrašus, pavyzdžiui VOIP ir vaizdo konferencijos, naujienų ir pašto siuntimas, programinės įrangos parsiuntimo, spausdinimo, nuotolinių prisijungimų ir kitokie leistinos veiklos įrašai. Taip pat reikalingi ir nepageidaujamos, ribojamos elgsenos įrašai.

7. Užpuolimų įvairovė ir naujumas: galimybė išanalizuoti įvairias ir naujas atakas bei grėsmės vystymosi laike scenarijus yra svarbi.

8. Anonimiškumas: reaguodami į privatumo teisinius reikalavimus, dauguma iš viešų tinklų rinktų duomenų rinkinių, negalėdami tinkamai nuasmeninti subjektų veiklos. visiškai pašalino paketų turinį, o tai sumažina duomenų rinkinio naudingumą, ypač kai kuriems aptikimo mechanizmams, tokiems kaip gilus paketų tikrinimas (DPI). Todėl svarbu, kad simuliuotame rinkinyje būtų apkrovos duomenų, kurių privatumas nesvarbus.

9. Heterogeniškumas: reikia turėti duomenis iš skirtingų šaltinių, pvz., tinklo srautai, operacinių sistemų žurnalai ar tinklo įrangos žurnalai, atminties vaizdai.

10. Požymių susietumas: tyrimams naudinga, kai pateikiami susiejami to paties įvykio įvairių tipų šaltinių, tokių kaip įrenginio atminties vaizdas, tinklo srautas ir įrenginių žurnalai, duomenys.

11. Dokumentacija: tinkamų dokumentų trūkumas yra vienas iš pagrindinių šios srities turimų duomenų rinkinių klausimų. Daugumoje duomenų rinkinių nėra dokumentų arba net jei jie ir yra, jie nėra išsamūs. Informacijos apie tinklo konfigūraciją, užpuolikų ir aukų mašinų operacines sistemas, atakos scenarijus stoka apsunkina tyrėjų darbą.

Toliau aptarti autoriaus tyrimuose panaudoti duomenys.

2.2.1 NSL-KDD

Viena iš pirmųjų mokslinių tyrimų reikmėms panaudotų kibernetinių įvykių duomenų rinkinių yra įsilaužimo aptikimo duomenų rinkinys „KDD'99“, sukurtas 1999 m. KDD taurės konkursui. Jis buvo atnaujintas ir dabar pateikiamas atsisiųsti New Brunswick universitete, Kanadoje, pavadinimu NSL – KDD. Šį rinkinį sudaro 41 matmuo (Bouzida & Cuppens, 2004).

Šis duomenų šaltinis buvo panaudotas pagrindinių komponenčių tyrimuose.

2.2.2 VU tinklo duomenys

Be išorinių šaltinių, atliekant šį tyrimą autorius rinko tinklo srautų įrašus Vilniaus universiteto tinkle. Šie pilno turinio duomenys transformuoti į formatą, leidžiantį juos toliau atvaizduoti N.Reddy et al. (MIT) disertacijoje aprašytais metodais.

2.2.3 CIC- IDS-2017

Mašininio mokymo algoritmų efektyvumo tyrimams naudotas atviros prieigos kibernetinės saugos tinklo duomenų šaltinis CIC IDS 2017, kuriame per atakų emuliavimo savaitę 2017 metų liepos 3-7 dienomis sukaupti duomenys. Tyrime naudoti MachineLearningCSV rinkinio failai, kuriuose pateikiami 71 tinklo parametras ir agregatai. Rinkmenų sąrašas pateikiamas 1 lentelėje:

Lentelė 1: Duomenų išrankų rinkmenos

Nr.	**Rinkmenos pavadinimas**
1	'Monday-WorkingHours.pcap_ISCX.csv'
.2	'Tuesday-WorkingHours.pcap_ISCX.csv'
3.	'Wednesday-workingHours.pcap_ISCX.csv'
4.	'Thursday-WorkingHours-Morning-WebAttacks.pcap_ISCX.csv'
5.	'Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv'
6.	'Friday-WorkingHours-Morning.pcap_ISCX.csv'
7.	'Friday-WorkingHours-Afternoon-PortScan.pcap_ISCX.csv'
8.	'Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv'

Tinklo srautų agregatų (NetFlow, CIC-FlowMeter) formatai yra išvestiniai, nuasmeninti, todėl dalis informacijos prarandama. Tačiau siekiant palengvinti duomenų analizę, šiuose duomenų formatuose pateikiamos statistinės metrikos (Count, Min, Max, Mean, Std, Len).

2.2.4 CIC-IDS-2018

Laikinių eilučių ir tyrimams naudotas atviros prieigos kibernetinės saugos tinklo duomenų šaltinis CIC IDS 2018, PCAP formato duomenys (lentelė 1 ir 2), kuriame sukaupti 80 tinklo parametrų.

Lentelė 2: Pilnų duomenų srautų rinkmenos

Name	Date modified	Type	Size
All 2018-02-23 between 14 and 17 UTC 172.31.69.20.pcap	2019-05-06 13:18	Wireshark capture...	66 679 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.21.pcap	2019-05-06 13:19	Wireshark capture...	310 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.22.pcap	2019-05-06 09:53	Wireshark capture...	497 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.23.pcap	2019-05-06 13:21	Wireshark capture...	52 545 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.24.pcap	2019-05-06 13:32	Wireshark capture...	63 900 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.25.pcap	2019-05-06 09:52	Wireshark capture...	350 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.26.pcap	2019-05-06 13:23	Wireshark capture...	26 427 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.28.pcap	2019-05-06 09:58	Wireshark capture...	10 956 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.29.pcap	2019-05-06 13:28	Wireshark capture...	56 472 KB
All 2018-02-23 between 14 and 17 UTC 172.31.69.30.pcap	2019-05-06 13:29	Wireshark capture...	60 343 KB

Lentelė 3: Pradiniai srauto (PCAP) duomenys

Ei. Nr.	**IP**	**Dydis, KB**	**Išranka, KB**	**Pastabos**
0	172.31.69.20	91’188	66’679	Normalus srautas
1	172.31.69.21	318	310	Normalus srautas
2	172.31.69.22	2’327	497	Normalus srautas
3	172.31.69.23	92’204	52’545	Normalus srautas
4	172.31.69.24	84’267	63’900	Normalus srautas
5	172.31.69.25	2’380	350	Normalus srautas, nepakanka duomenų
6	172.31.69.26	46’350	26’427	Normalus srautas
7	172.31.69.28	13’714	10’956	Normalus ir anomalus srautas, atrinktas tyrimo objektas su įsilaužimu
8	172.31.69.29	82’331	56’472	Normalus ir anomalus srautas, atrinktas tyrimo objektas su įsilaužimu
9	172.31.69.30	86’750	60’343	Normalus srautas

Fraktalinės dimensijos tyrime panaudoti lentelėje 4 pateikti duomenys::

Lentelė 4: CIC-IDS-2018 duomenų bazės IP 172.31.69.28 duomenų rinkinys

Ei. Nr.	Duomenų šaltinis	Dydis, KB	Išranka[4], KB
1	„Friday-23-02-18_TrafficForML_CICFlowMeter.csv“	374’892	693

[4] Pastaba: Tyrime naudota 1/1000 normalaus srauto ir visas anomalus srautas

3 Duomenų paruošimo aptarimas

3.1 Duomenų valymas ir atliktos transformacijos

Šiame tyrime įvairias metodais nagrinėjami CIC-IDS-2017 rinkinyje surinktų 14 skirtingų atakų rūšių duomenys pateikiami 5 lentelėje:

Lentelė 5: CIC-IDS-2017 sudėtis

Klasė	**Įrašų skaičius**	**Pašalinta**	**Dalis**
BENIGN	2273097	176613	7,77%
DoS Hulk	231073	58224	25,20%
PortScan	158930	68111	42,86%
DDoS	128027	11	0,01%
DoS GoldenEye	10293	7	0,07%
FTP-Patator	7938	2005	25,26%
SSH-Patator	5897	2678	45,41%
DoS slowloris	5796	411	7,09%
DoS Slowhttptest	5499	271	4,93%
Bot	1966	13	0,66%
Web Attack - Brute Force	1507	37	2,46%

Importuojant įkelti 2830743 įrašai, 79 stulpeliai. Objektinio tipo įrašams pritaikyti duomenų tipo konverteriai. Pašalinti 1358 nuliniai įrašai. Pašalinti dubliuoti įrašai (paliekant pirmąjį). Toks įrašų dublikatų šalinimas realaus laiko sistemose realizuojamas skaitliukų būdu. Pašalinti 8 tušti požymiai (std == 0). Realiam sraute negalėtume žinoti, kad jie nebus populiuoti duomenimis. 1211 neuždarytuose flows įrašuose begalybės pakeistos klasės maksimumais (realiame sraute galime kaupti stebimas didžiausias reikšmes, bet jos gali keistis; klausimas kodėl jos šiame šaltinyje iš viso egzistuoja neatsakytas. Galimai tai likę neuždaryti per 5 minučių srauto stebėjimo laiką srautai (flows)). Normalizuotos reikšmės. Realiame tinklo sraute negalime žinoti tikėtino minimumo ir maksimumo, tačiau įrašus normalizuoti galime sukaupę pakankamai duomenų. Šiuo atveju rinkinyje pateikiami vienos savaitės duomenys.

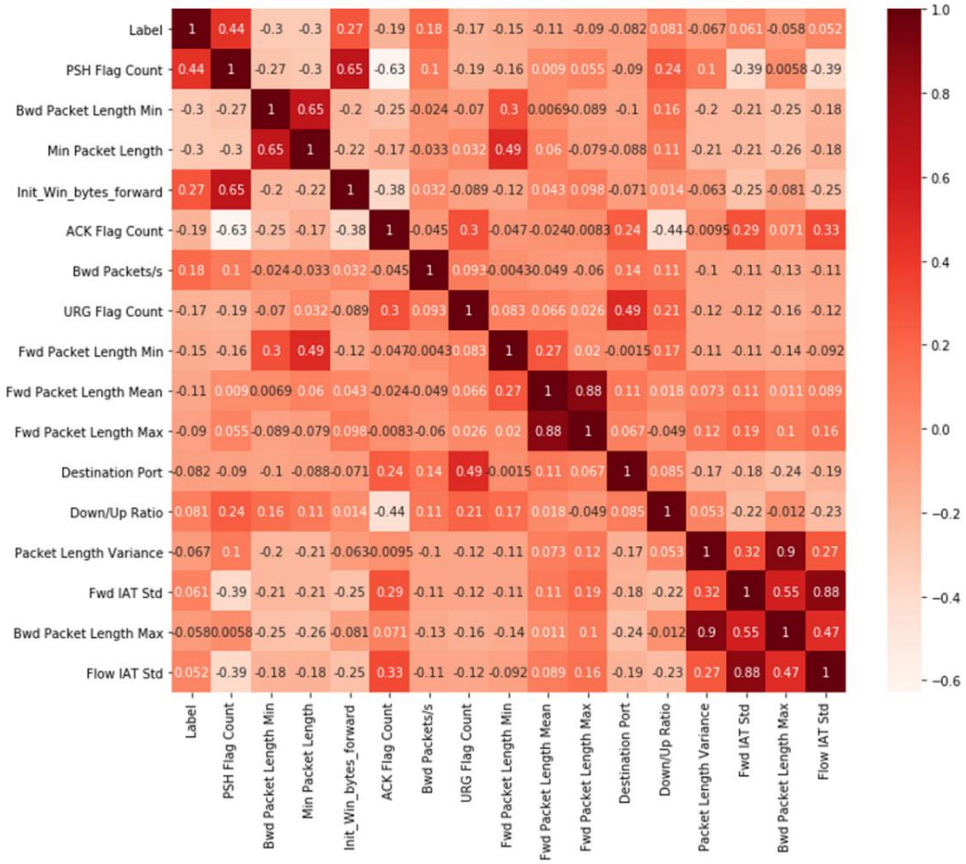
Nustačius, kad pagrindinė duomenų klasė yra atstovaujama neproporcingai, atsitiktine tvarka pašalinta tiek įrašų, kad likę kibernetinių atakų duomenys sudarytų 50% visų duomenų. Siekiant išvengti mažoms klasėms būdingo nepatekimo į stratifikuotas atrankas, mažiausiai įrašų turinčios klasės praturtintos iki 30 įrašų.

3.2 Duomenų šaltinio dedamųjų tyrimas

Šiame etape duomenys buvo nagrinėti siekiant išsiaiškinti svarbiausius požymius, įnešančius didžiausią indėlį į mašininio mokymo tikslumą. Papildant ankstesnių metų tyrimus, kuriuose požymiai nagrinėti pagrindinių komponenčių (PCM) ir Šapo verčių nustatymo metodais, išbandyti dar trys skirtingi būdai parenkant savybes: koreliacijų analizė, reikšmingiausių savybių atrinkimas KBest metodu ir rekursinis reikšmingiausių savybių atrinkimas (RFE).

3.2.1 Dedamųjų filtravimas

Filtruojant įrašus atlikti šie veiksmai: a) nustatyta koreliacija su prognozuojama verte, pašalintos nereikšmingos (<0,05) savybės, b) nustatyta savybių tarpusavio koreliacija, pašalintos perteklinės savybės (>0,90). Rezultatas – 16 atrinktų savybių. Sprendimo kiek palikti savybių ribos nustatymas toliau galėtų būti vykdomas pagal norimą prognozės tikslumą. Toliau 1 paveiksle pateikiamas atrankos rezultatas:



pav. 1 Koreliacijos analizės rezultatai

3.2.2 Reikšmingiausių dedamųjų atrinkimas (KBest)

Pasirinktas Python Scikit-Learn bibliotekos metodas SelectKBest, atrinktos 20 dedamųjų:

lentelė 5: Kbest metodu atrinktos 20 dedamųjų

Nr.	Pavadinimas
1	Flow IAT Max
2	Max Packet Length
3	Fwd IAT Std
4	Avg Bwd Segment Size
5	Packet Length Std
6	Fwd IAT Total
7	Fwd IAT Max
8	Idle Min
9	Flow Duration
10	Packet Length Mean
11	Active Mean
12	Idle Mean
13	Bwd Packet Length Max
14	Active Min
15	Flow IAT Std
16	Bwd Packet Length Std
17	Idle Max
18	Average Packet Size
19	Bwd Packet Length Mean
20	PSH Flag Count

SelectKBest tiesiog įvertina dedamąsias naudodama f_classif funkciją (kuri apskaičiuoja pateiktos dedamosios ANOVA F vertę, bet gali būti ir kitos), tada pašalina visas, išskyrus k, aukščiausius balus gavusias dedamąsias.

3.2.3 Rekursinis reikšmingiausių dedamųjų atrinkimas (RFE)

Pasirinktas Scikit-Learn bibliotekos metodas RFE (Recursive Feature Elimination) – remiantis kitu apmokytu modeliu po vieną eliminuojamos duomenų savybės (Guyon et al., 2002). Šiuo atveju pasirinktas greitai besimokantis sprendimų medis: RFECV(estimator=cart, scoring='f1_weighted,...). Atrinktos 16 svarbiausių savybių:

lentelė 6 Rekursijos metodu atrinktų savybių sąrašas

Nr.	Pavadinimas
1	Fwd IAT Mean

2	Total Length of Fwd Packets
3	Avg Fwd Segment Size
4	Avg Bwd Segment Size
5	Destination Port
6	Subflow Fwd Packets
7	Packet Length Std
8	Fwd IAT Total
9	Fwd IAT Min
10	Init_Win_bytes_forward
11	Fwd Packet Length Max
12	Total Fwd Packets
13	Flow IAT Std
14	Bwd Packet Length Std
15	Init_Win_bytes_backward
16	Bwd Packet Length Mean

3.3 Reikšmingiausių dedamųjų atrinkimo rezultatai

Panaudojus filtravimo metodą, duomenų šaltinyje nustatyta tiek nereikšmingų tiek stipriai koreliuotų savybių. Tai vėliau patvirtina ir QDA analizė.

Su 20 reikšmingų savybių pasiekiamas duomenų šaltinio autorių publikacijoje nurodytas tikslumas.

Geriausias tikslumas atrenkant savybes (paliekant tą patį savybių skaičių skirtingais metodais) pasiektas panaudojat Kbest.

3.4 Duomenų dedamųjų generavimas

Trečiame tyrimo etape, papildant duomenų šaltinių agregatų sąrašą, pasiūlytas papildomas statistinis agregatas, fraktalinė dimensija. Fraktalinė dimensija gali būti interpretuojama kaip vaizdo sudėtingumo matas.

Fraktalinės dimensijos apibrėžimą pirmą kartą pasiūlė Benoitas Mandelbrotas (Mandelbrot, 1983).

Tegul $H(R_m)$ yra visų Euklido erdvės R_m visų nenuoseklių kompaktiškų pogrupių erdvė. Teigiama, kad kompaktiškas rinkinys A R_m yra panašus į save, jei A yra skirtingų r spindulio skritulių, reikalingų A padengti, sąjunga.

Kad būtų paprasčiau, kaip pasiūlė Higuchi (Higuchi, 1988), R_m padengtas kvadratiniais langeliais, kurių šoniniai ilgiai r.

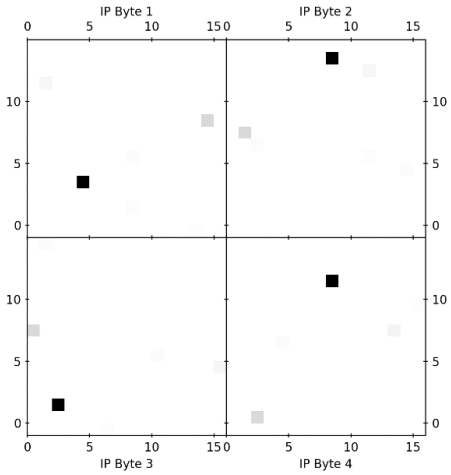
Kiekvienam r> 0 tegul $N_r(A)$ yra mažiausias langelių, esančių r šone, susikertančių A, skaičius.

Jei $D_f(A) = \lim_{r \to 0} \frac{\ln N_r(A)}{\ln\left(\frac{1}{r}\right)}$ egzistuoja, tada $D_f(A)$ vadinamas A fraktalo matmeniu arba langelio matmeniu.

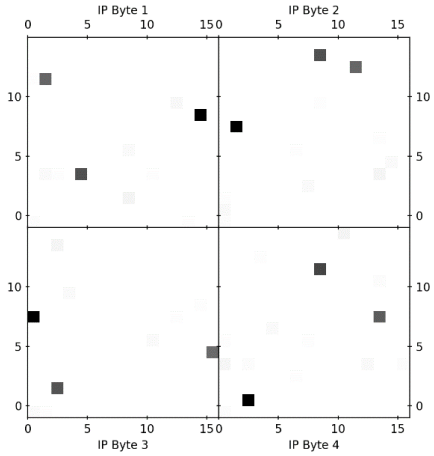
3.5 DOS ir DDOS analizė fraktalinės dimensijos metodu

Šioje eksperimento dalyje galimam kibernetiniam išpuoliui nustatyti naudojamas T. Higuchi Boxcount algoritmas (Higuchi, 1988). Šis DOS ir DDOS atakų atpažinimo algoritmas pasiūlytas (Xia et al., 2012) tinka ir Brute Force atakoms.

Tinklo srauto Hausdorfo[5] fraktalinė dimensija (Hausdorff, 1918) paskaičiuojama Higuchi metodu (Higuchi, 1988). Šiuo metodu aproksimuojamas Kim, Reddy ir Vannucci (Kim et al., 2004) metodu parengto vaizdo (pav. 2a ir 2b), sudėtingumas.



Pav. 2 a. DOS atakos pradžioje



Pav. 2b. DOS atakai prasidėjus

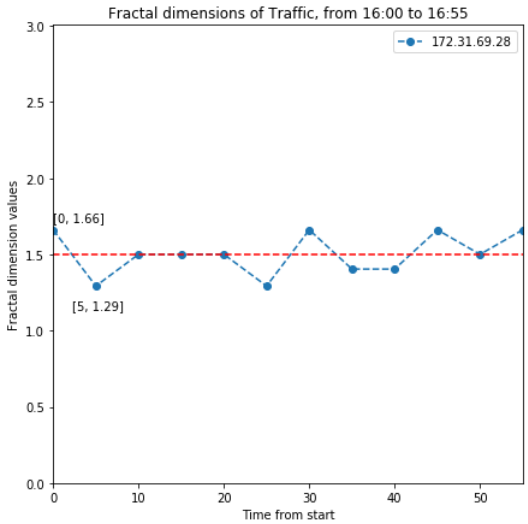
Fraktalinės dimensijos rodiklis toliau naudojamas kaip indikatorius aptinkant kibernetinį DOS incidentą. Algoritmas perrenka visus vienspalvio vaizdo Z taškus kurių yra N ($m*n = c*n^2$, kur m ir n vaizdo išmatavimai pikseliais, o c – skalės koeficientas, pvz. 4:3, 16:9), poromis (kiekvieną su kiekvienu), atlieka konstantos tikslumu panašų skaičių aritmetinių operacijų, todėl algoritmo sudėtingumas yra proporcingas taškų porų skaičiui. Tačiau, įvertinus taškų

[5] Felix Hausdorff (1868 –1942), vokiečių matematikas

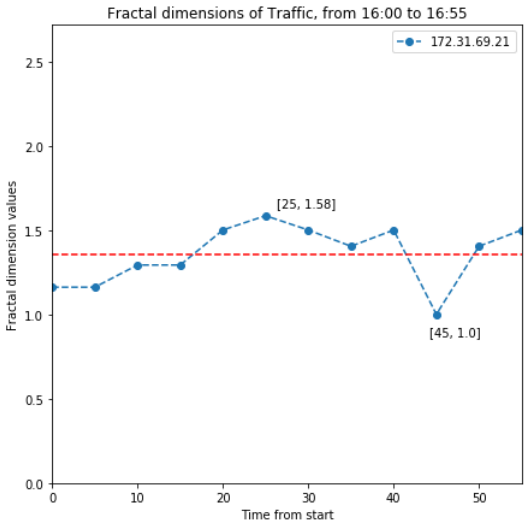
retumą, galimas algoritmo spartinimas, praleidžiant nulines reikšmes turinčias taškų poras, o tada griežtas asimptotinis sudėtingumo vertinimas būtų $\Theta(n^k)$ kur $k \in [\log_2 3, 2]$. Čia $O(n^{\log_2 3})$ yra daugybos Karatsuba algoritmu sudėtingumas, o $O(n^2)$ paprastos daugybos stulpeliu sudėtingumas.

Tinklo duomenų srautas eksperimento metu buvo dalinamas į vienodos trukmės rinkmenas ir suskaičiuotos ryšio seansų poros tarp įvairių išorinių kompiuterių ir tiriamojo. Tuomet sugrupavus duomenis, paskaičiuotos srautų kiekių reikšmės laiko intervale ir Box-counting metodu suskaičiuota trečio IP adreso okteto duomenų, sugrupuotų į vaizdą 16x16, fraktalinė dimensija. Eksperimentas atliktas su 30s, 1, 2 ir 5 minučių intervalais, perrenkant 10 tiriamųjų IP adresų iš CIC-IDS-2018 duomenų bazės.

Fraktalinės dimensijos kaita atvaizduojama Pav. 3a ir Pav. 3b. parodo DOS ataką tais atvejais, kai dėl kompiuterio perkrovimo duomenų srautas nutrūksta. Tokia situacija Pav. 3b. stebima 16:45, kai fraktalinės dimensijos reikšmė krinta iki mažiausios galimos pagal Higuchi aproksimuojantį algoritmą d = 1. O kai vaizdo užpildymas didėja, dimensijos reikšmė galimai artėja link d = 2.



Pav. 3 a. Įprastinės veiklos fraktalinės dimensijos dinamika



Pav. 3 b. Stebimas minimumas Per ataką 16:45

Padaryta prielaida, kad fraktalinės dimensijos pokytis indikuos ataką. Tyrimas parodė, kad nors duomenų statistinis pasiskirstymas turi įtakos fraktalinės dimensijos absoliutinei reikšmei, dimensijos reikšmės nuokrypis nuo vidurkio yra akivaizdus tuomet, kai užpuolikui laužiantis generuojami dideli neįprastų naujų įvykių srautai ir dimensijos reikšmė reikšmingai

pakyla virš normalaus veiklos srauto fraktalinės dimensijos vidurkio (d=1.5). Tuo tarpu, įvykus įsilaužimui, fraktalinė dimensija dėl taškų porų skaičiaus sumažėjimo krenta ir rodo artimas d= 1 reikšmes, jei parinktas ilgesnis agregavimo periodas (2, 5 minutės). Tuo būdu, kai fraktalinės dimensijos reikšmė stebimu periodu statistiškai reikšmingai išauga virš ilgalaikio stebėjimo laikotarpio vidurkio (1 val.) ir išvestinis pokytis rodo reikšmingą skirtumą su ankstesniu periodu (30s, 1 min, 2 min, 5 min), galima prognozuoti įsilaužimą. Kai įsilaužimas (pagal faktinius duomenis) jau yra įvykęs, duomenų srautas ženkliai sumažėja, dėl ko tiek statistiniai tiek fraktalinės dimensijos skaičiavimo rezultatai indikuoja minimalias reikšmes (d = 1). Skaičiuojant 30 sekundžių intervale minimumas akivaizdus, tačiau skaičiavimai ir algoritmas darosi nestabilūs dėl reikšmių trūkumo. Šio trūkumo lengvai pavyksta išvengti papildant duomenis standartinio nuokrypio (Z-score) skaičiavimais, kurie yra pakankami ir veikia su mažesniu taškų skaičiumi.

4 Mašininio mokymo algoritmų taikymas

Python aplinkoje su CIC-IDS-2017 duomenimis realizuoti 7 lyginamojo duomenų šaltinio CIC-IDS-2017 autorių publikuoto tyrimo mašinų mokymo algoritmai (KNN, RFC, CART, Adaboost, Gaussian Naive Bayes, QDA, MLP), ir palygintas gautas efektyvumas.

Papildomai tam pačiam uždaviniui Python Keras –Tensorflow 2.1 aplinkoje realizuoti ANN ir CNN neuroniniai tinklai.

Siekiant pagerinti rezultatus, panaudojant tinklelio paieškos (‚Grid Search')metodą atrinkti skirtingų algoritmų mokymo hiperparametrai.

Suskaičiuoti algoritmų tikslumo (accuracy), klasių prognozės tikslumo (precision), jautrumo (recall), harmoninio tikslumo F1 (7 lentelė), Hamming Loss ir Jaccart Score, subalansuoto tikslumo (8 lentelė, 4 paveiksas) rodikliai.

lentelė 7 Mašininių algoritmų sulyginimas

Algoritmas	**Tikslumas**		**Jautrumas**		**F1**		**Trukmė (s)**	
	PUB	EKSP	PUB	EKSP	PUB	EKSP	PUB	EKSP
KNN	0,96	0,989	0,96	0,989	0,96	0,985	1908	897
RFC	0,98	0,991	0,97	0,993	0,97	0,992	74	36
ID3 vs. CART	0,98	0,997	0,98	0,998	0,98	0,998	235	7

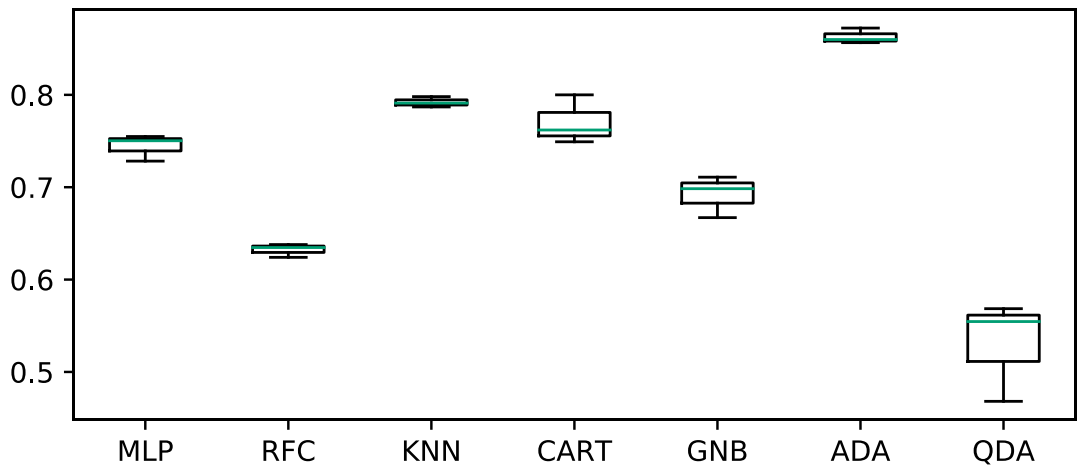
Adaboost	0,77	0,999	0,84	0,999	0,77	0,999	1126	278
Naive-Bayes	0,88	0,82	0,04	0,87	0,04	0,84	15	4
ANN	-	0,986	-	0,986	-	0,796	-	294
CNN	-	0,994	-	0,994	-	0,865	-	320

lentelė 8 Mašininių algoritmų sulyginimas

Algoritmas	**Tikslumas**	**BAS**	**Hamming**	**Jackart**
KNN	0,988	0,792	0,002	0,978
RFC	0,987	0,632	0,011	0,975
CART	0,997	0,770	0,003	0,989
Adaboost	0,998	0,862	0,001	0,995
Naive-Bayes	0,789	0,692	0,210	0,734
ANN	0,986		0,013	0,987
CNN	0,997		0,002	0,992
MLP	0,984	0,744	0,015	0,970
QDA	0,839	0,530	0,161	0,726

Subalansuoto tikslumo rodiklis BAS, kuris skaičiuojamas kaip klasių tikslumo svertinis vidurkis geriausiai iš skaičiuotų rodiklių parodo algoritmų skaičiavimo efektyvumą.

Pav. 4 Subalansuoto tikslumo rodiklis



Rezultatai patikrinti panaudojus kryžminio patikrinimo ('cross-validation') ir validavimo rinkinio atidėjimo metodus.

Nagrinėjant Adaboost ansamblio, eksperimente sudaryto iš 100 CART sprendimų medžių, kryžminio patvirtinimo matricą (9 lentelė) nustatyta, kad ansamblis atsižvelgė ir į mažai reprezentuotas klases, nes aukšti ir kiti klaidą arba šiuo atveju tikslumą parodantys rodikliai.

Klasė	**Tikslumas**	**Jautrumas**	**F1**	**Reikšmių skaičius**
BENIGN	0,998	0,999	0,999	85175
Bot	0,916	0,839	0,876	391
DDoS	1	1	1	25603
DoS GoldenEye	0,998	0,993	0,996	2057
DoS Hulk	0,999	0,998	0,999	34569
DoS Slowhttptest	0,995	0,991	0,993	1046
DoS slowloris	0,996	0,992	0,994	1077
FTP-Patator	1	0,998	0,999	1187
Heartbleed	1	1	1	6

5 Išvados ir rezultatai

Panaudojus hiper-parametrus, atrinktus GridSearch pagalba, panaudojant klasikinius mašininio mokymo metodus pavyko pasiekti geresnių rezultatų, nei duomenų šaltinio autorių publikacijoje.

Papildomai sumodeliuotų ANN ir CNN rezultatai geresni, nei duomenų šaltinio autorių gautieji klasikiniais algoritmais, bet blogesni, nei šiame tyrime atliktų skaičiavimų.

Vertinant subalansuoto tikslumo rodiklį galima teigti, kad praktiniams įsilaužimo nustatymo uždaviniams tiksliausiais Adaboost ansamblis.

6 Literatūros sąrašas

Bouzida, Y., & Cuppens, F. (2004). Efficient intrusion detection using principal component analysis. *Proceedings of The*. http://yacine.bouzida.free.fr/Articles/2004SAR.pdf

Gharib, A., Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2016). An Evaluation Framework for Intrusion Detection Dataset. *2016 International Conference on Information Science and Security (ICISS)*, 1–6. https://doi.org/10.1109/ICISSEC.2016.7885840

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*. https://doi.org/10.1023/A:1012487302797

Hausdorff, F. (1918). Dimension and outer measure. *Mathematische Annalen*, *79*(1–2), 157–179. https://doi.org/10.1007/BF01457179

Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*, *31*(2), 277–283. https://doi.org/10.1016/0167-2789(88)90081-4

KDD. (1999). *KDD Cup 1999 Data*. KDD Cup 1999 Data. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

Kim, S. S., Reddy, A. L. N., & Vannucci, M. (2004). *Detecting Traffic Anomalies through Aggregate Analysis of Packet Header Data*. *June*, 1047–1059. https://doi.org/10.1007/978-3-540-24693-0_86

Laboratory, L. (1998). *Darpa intrusion detection data sets (1998)*. Http://Www.Ll.Mit.Edu/r-d/Datasets/1998-Darpa-Intrusion-Detection-Evaluation-Dataset. http://www.ll.mit.edu/r-d/datasets/1998-darpa-intrusion-detection-evaluation-dataset

Lawrence Berkeley National Laboratory. (2010). *The Internet Traffic Archive*. http://ita.ee.lbl.gov/index.html

Mandelbrot, B. B. (1983). *The fractal geometry of nature*. Henry Holt and Company.

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, *January*, 108–116. https://doi.org/10.5220/0006639801080116

The Cooperative Association for Internet Data Analysis. (2010). CAIDA - The Cooperative Association for Internet Data Analysis. In *CAIDA*. http://www.caida.org/home/

The Shmoo Group. (2011). *Defcon*.

Xia, Z., Lu, S., & Li, J. (2012). DDoS flood attack detection based on fractal parameters. *2012 International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2012*, *2*, 1–5. https://doi.org/10.1109/WiCOM.2012.6478475