VILNIUS TECH
Vilnius Gediminas Technical University

Mokslas – Lietuvos ateitis / Science – future of Lithuania
ISSN 2029-2341 / eISSN 2029-2252
2023 Volume 15, Article ID: mla.2023.19385, 1–9
https://doi.org/10.3846/mla.2023.19385

Information technologies and multimedia
Informacinės technologijos ir multimedija

ĮSILAUŽIMŲ APTIKIMAS KOMPIUTERIŲ TINKLUOSE TAIKANT HIBRIDINIUS MAŠININIO MOKYMOSI METODUS

Karina ČIURLIENĖ *, Denisas STANKEVIČIUS

Vilniaus Gedimino technikos universitetas, Vilnius, Lietuva

Gauta 2023 m. birželio 15 d.; priimta 2023 m. birželio 22 d.

Santrauka. Viena iš aktualių kibernetinės saugos tyrimų krypčių – tai įsilaužimų arba anomalijų aptikimas kompiuterių tinkle. Įsilaužimų skaičius nuolat didėja, o taikomos įsilaužimo technikos ir metodai sudėtingėja, todėl siekiant apsaugoti kompiuterių tinklą, reikia taikyti vis sudėtingesnius apsaugos metodus. Tinklo įsilaužimams ir anomalijoms nustatyti taikomi įvairūs mašininio mokymosi algoritmai, tačiau jų tikslumas yra ribotas. Siekiama pagerinti tinklo anomalijų aptikimą, taikomi hibridiniai mašininio mokymosi algoritmai. Straipsnyje pasiūlyti trys nauji hibridiniai mašininio mokymosi algoritmai, ištirtas jų tikslumas naudojant du viešai prieinamus duomenų rinkinius, t. y. CSE-CIC-IDS2018 ir NSW-NB-15. Siekiant padidinti klasifikavimo modelių tikslumą, buvo atliktas hiperparametrų optimizavimas. Reikšmingiems duomenų rinkinio požymiams nustatyti taikytas iteracijų metodas ir Chi kvadrato χ^2 testas. Analizuojant tyrimo rezultatus, nustatyta, kad aukščiausias tinklo anomalijų atpažinimo tikslumas 99,34 % buvo pasiektas taikant hibridinį algoritmą, sudarytą iš sprendimų medžio, naivaus Bajeso ir daugiasluoksnio perceptrono algoritmų rinkinio. Šis rezultatas yra 3,13 % geresnis, lyginant su geriausiu tikslumu, gautu taikant atskirus mašininio mokymosi algoritmus. Siekiant kompleksiškai įvertinti tirtus mašininio mokymosi algoritmus ir jų tinkamumą įsilaužimams kompiuterių tinkle aptikti, algoritmai buvo sureitinguoti taikant SCR, DR, FR reitingavimo metodus.

Reikšminiai žodžiai: tinklo anomalijos, mašininis mokymasis, χ2 Chi kvadratu testas, hiperparametrai, hibridiniai algoritmai.

Įvadas

Šiuolaikiniame pasaulyje sparčiai vystantis informacinėms technologijoms, debesų kompiuterijos sistemoms, vis daugiau paslaugų skaitmenizuojama. Jau tapo įprasta naudotis e. valdžios ar e. komercijos paslaugomis, pasiekiamomis internetu iš bet kurios pasaulio vietos (Aminanto & Kim, 2016). Toks platus paslaugų naudojimas internete siejamas ir su kibernetinėmis grėsmėmis, kylančiomis dėl nepakankamos duomenų apsaugos, blogai tvarkomų konfigūracijų, žemo vartotojų žinių lygio. Kibernetiniai nusikaltėliai įvairiais būdais bando įsilaužti, blokuoti prieigas prie skaitmeninių paslaugų, trikdo jų veiklą, naudoja užkrėstus kompiuterius kibernetinėms atakoms, o pritaikydami socialinės inžinerijos metodus, išgauna konfidencialius vartotojo duomenis (Ahmed et al., 2016). Remiantis kibernetinės saugos ataskaitomis, matyti, kad daugiau nei pusė įmonių per pastaruosius metus patyrė bent vieną saugos incidentą, o atakų metu padarytų žalų suma siekia iki 6 trilijonų JAV dolerių (Trustware, 2020). Pagrindiniai saugos incidentai susiję su duomenų nutekinimu, tapatybės duomenų praradimu, sukčiavimu ir įsibrovimu į tinklą. Kompanijos *Trustware* atliktas tyrimas nustatė, kad įsilaužimas į tinklą nustatomas vidutiniškai per 11 dienų (Trustware, 2020). Tai reiškia, kad visą tą laiką nusikaltėlis rinko duomenis, stebėjo tinklo infrastruktūrą, analizavo sistemų pažeidžiamumus. Siekiant valdyti ir mažinti įsilaužimus į tinklą, naudojamos įsilaužimų aptikimo sistemos (Bao et al., 2019). Šios sistemos skenuoja tinklo paketus ir klasifikuoja juos atskirdamos gerus ir susikompromitavusius. Tačiau, kaip rodo praktika, ne visi paketai klasifikuojami teisingai, todėl pagrindinis uždavinys, susijęs su šių sistemų tobulinimu, yra jų tikslumo ir stabilumo didinimas (Ashiku & Dagli, 2021).

Įsilaužimams aptikti taikomi parašu, anomalijomis arba protokolo būsena pagrįsti metodai. Parašu paremtas aptikimas analizuoja tinklo srautą lygindamas skenuojamus objektų parašus su kenkėjiškais ir tokiu būdu vykdo klasifikavimą (Bhuyan et al., 2014). Tokių sistemų

*Autorius susirašinėti. El. paštas *karina.ciurliene@vilniustech.lt*

trūkumas yra tas, kad anomalijos nustatomos tik tuo atveju, jei sistema turi atitinkamą parašą, o tai reiškia, kad nežinomos atakos yra praleidžiamos. Anomalijomis pagrįstais aptikimo metodais tiriamas tinklo srautas taikant mašininio mokymosi metodus nuokrypiams nustatyti (Kwon et al., 2019). Tikimybės reikšmei viršijus nustatytą slenkstį, įvykis laikomas anomaliu ir siunčiamas įspėjimas administratoriui. Taikant tokius metodus galima aptikti naujas anomalijas, tačiau reikia didesnių skaičiavimo resursų, o jų tikslumas priklauso nuo modelio ir apmokymui skirtų duomenų rinkinio dydžio (Ferrag et al., 2020). Protokolo būsenos analize paremti metodai lygina žinomus protokolų profilius su tinklo srautu, naudodami iš anksto nustatytus, tiekėjo pateiktus profilius, kad nustatytų atsitiktinę komandų seką tinklo ir taikomųjų programų lygmenyse (Ashiku & Dagli, 2021).

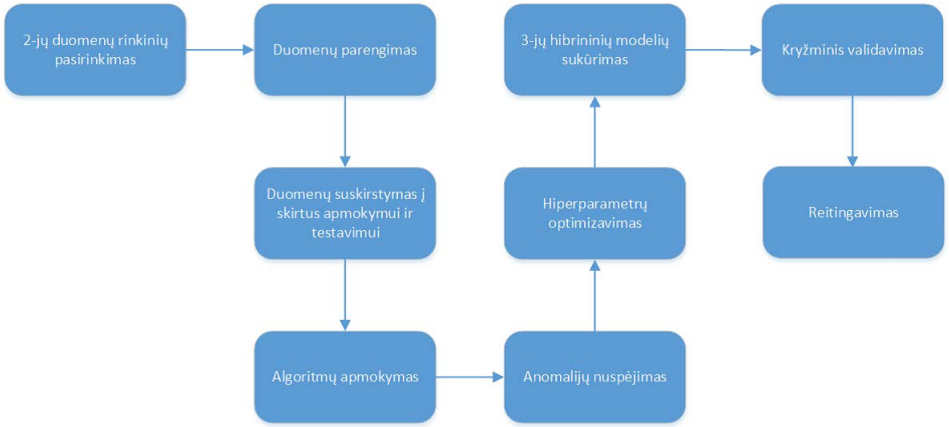
Visi apžvelgti įsilaužimo aptikimo metodai turi savų privalumų ir trūkumų, tačiau anomalijomis pagrįsti metodai leidžia aptikti nežinomas, nulinės dienos atakas, tas suteikia privalumą, lyginant su kitais aptikimo metodais (Kwon et al., 2019). Siekiant padidinti šių metodų tikslumą, taikomi sprendimai, susiję su duomenų rinkinio kokybės gerinimu arba modelio struktūros tobulinimu (Kanimozhi & Jacob, 2019). Duomenų rinkiniui tvarkyti naudojamos tokios technikos kaip svarbiausių atributų analizė ir atrinkimas, duomenų valymas, balansavimas, vektorizavimas, normalizavimas, duomenų rinkinio sintetinis didinimas. Tokios anomalijų aptikimo sistemos tobulinimo kryptys leidžia pasiekti didesnį tikslumą, tačiau paprastai didina modelio apmokymo laiką, todėl praktinis tokių sprendimų pritaikymas įsilaužimų aptikimo sistemose yra probleminis, ypač kai kalbame apie greitą modelio atnaujinimą, siekiant aptikti naujus įsilaužimų tipus ar klases (Atefinia & Ahmadi, 2021; Bulavas et al., 2021). Kita aptikimo sistemos tobulinimo kryptis siejama su mašininio mokymosi algoritmų tobulinimu, sujungimu, hiperparametrų optimizavimu, hibridinių metodų taikymu. Šios technikos leidžia didinti tikslumą, mažinti modelio dydį, procesoriaus apkrovą, užimamos atminties dydį.

Megantara ir Ahmad (2021) pasiūlė hibridinį mašininį mokymosi metodą įsilaužimams aptikti, paremtą svarbiausių duomenų atributų parinkimu ir duomenų mažinimu. Atlikus eksperimentus su UNSW-NB15 duomenų rinkiniu buvo nustatyta, kad galima pagerinti klasifikavimo tikslumą iki 91,86 %. Hassan et al. (2020) tyrė hibridinį giliojo mokymosi modelį, skirtą tinklo įsilaužimams aptikti. Autoriai sudarė konvoliucinį neuroninį tinklą (CNN) ir sumažino svorius, naudodami ilgalaikės trumpalaikės atminties (WDLSTM) neuroninį tinklą. Konvoliucinis neuroninis tinklas leido rasti reikšmingus požymius, o WDLSTM buvo panaudotas, kad būtų išlaikyta ilgalaikė išskirtų požymių priklausomybė ir išvengta persimokymo rekurentiniuose ryšiuose. Chkirbene et al. (2020) pasiūlė hibridinį dviejų mašininio mokymosi algoritmų derinį, kuriame naudojo atsitiktinio miško algoritmą reikšmingiausioms duomenų rinkinio savybėms rasti ir klasifikavimo bei regresijos medžių algoritmą (CART) atakoms klasifikuoti. Atlikus eksperimentą su UNSW-NB15 duomenų rinkiniu, buvo nustatyta, kad pasiūlytas metodas leidžia pagerinti klasifikavimo tikslumą.

Šiame darbe pasiūlyti trys nauji hibridiniai mašininio mokymosi metodai, atlikta jų analizė, hiperparametrų optimizavimas. Tyrimas atliktas su dviem viešai prieinamais duomenų rinkiniais, nustatyti reikšminiai rinkinio atributai. Tyrimas leido nustatyti ir sureitinguoti pasiūlytų hibridinių mašininio mokymo metodų tikslumą ir greitaveiką bei juos palyginti.

1. Tyrimų metodika

Pagrindinis tyrimų tikslas buvo pagerinti įsilaužimų aptikimą, pasiūlant patobulintus mašininio mokymosi metodus. Siekiant įgyvendinti išsikeltą tikslą buvo sudaryta tyrimų metodika, kuri apėmė tokius žingsnius: duomenų rinkinių parinkimą, duomenų paruošimą tyrimui, duomenų padalinimą į duomenis, skirtus apmokymui ir testavimui, mašininio mokymosi metodų tyrimui parinkimą, apmokymą ir testavimą, hiperparametrų optimizavimą. Remiantis gautais rezultatais buvo atrinkti mašininio



1 paveikslas. Tyrimų eigos schema
Figure 1. Research flowchart

mokymosi metodai, iš kurių sudaryti hibridiniai mašininio mokymosi metodai, atliktas jų apmokymas ir testavimas taikant kryžminio validavimo principą. Galiausiai pasiūlyti hibridiniai metodai buvo sureitinguoti taikant SCR, DR, FR reitingavimo metodus (1 pav.).

1.1. Duomenų rinkiniai ir paruošimas

Mašininio mokymosi metodų kokybei įvertinti naudojama sumaišymo matrica, kuri atvaizduoja algoritmo esminį efektyvumą ir parodo, koks yra teisingai ir neteisingai suklasifikuotų duomenų skaičius. Minėtų dydžių vertinimas be istorinio konteksto nėra labai efektyvus, todėl naudojamos ir išvestinės metrikos, tokios kaip tikslumas, jautrumas, preciziškumas, F1 metrika. Šių metrikų reikšmės apskaičiuojamos naudojant sumaišymo matricos duomenis (Mohri et al., 2018).

Tinklo įsilaužimų aptikimas naudojant mašininį mokymą yra paremtas duomenų analize, todėl reikia surinkti duomenis, kurie bus skirti modeliui apmokyti ir testuoti. Duomenys turi apibūdinti įprastinį tinklo darbą ir kenkėjiškas atakas. Prieš duomenis naudojant modelyje, reikia juos apdoroti, sužymėti ir tik tada galima duomenų rinkinį naudoti. Tyrime buvo naudojami du viešai prieinami duomenų rinkiniai, t. y. Kanados kibernetinio saugumo instituto sukurtas duomenų rinkinys CSE-CIC-IDS2018 ir Australijos kibernetinio saugumo centro parengtas duomenų rinkinys UNSW-NB-15.

CSE-CIC-IDS2018 rinkinyje surinkti duomenys apie normalų tinklo srautą ir 14 kibernetinių tinklo atakų, tokių kaip „Brute-force“, „Heartbleed“, „Botnet“, „DoS“, „DDoS, kuriuos apibūdina 80 atributų, gautų iš tinklo srauto paketų ir įvykių žurnalų. Duomenys rinkti iš 50 virtualių mašinų tinklo, kurį sudarė asmeniniai kompiuteriai ir serveriai. Rinkinį sudaro 10 failų CSV formatu, surinkti naudojant *pcap* įrankį. Bendras duomenų rinkinio dydis 6,41 GB. Prieš atliekant duomenų klasifikavimą, visi 10 failų buvo sujungti į vieną CSV failą. Atlikus duomenų rinkinio analizę, nustatyta, kad duomenų rinkinys nesubalansuotas, t. y. normalaus tinklo srauto įrašai sudaro 87,07 % visų įrašų, o atakų įrašų skaičius apima nuo 4,2 % iki 0,001 %.

UNSW-NB-15 rinkinį sudaro normalaus tinklo srauto ir 9 atakų, tokių kaip „Fuzzers“, „Analysis“, „Backdoors“, „DoS“, „Exploits“, „Generic“, „Reconnaissance“, „Shellcode“ ir „Worms“, duomenys, kur kiekvieną įrašą apibūdina 49 atributai. Visi atributai sugrupuoti į tokias klases: srauto atributai, turinio atributai, pagrindiniai atributai, papildomi atributai, klasifikavimo atributai. Duomenys sugeneruoti *tcpdump* ir *pcap* įrankiais, surinkta 100 GB duomenų. Kaip ir pirmasis duomenų rinkinys, UNSW-NB-15 rinkinys yra nesubalansuotas, t. y. normalaus tinklo srauto įrašai sudaro 67,84 % visų įrašų. Atakų įrašų skaičius apima nuo 13,52 % iki 0,07 %.

Abiejuose duomenų rinkiniuose dalis atributų yra tokie pat, t. y. protokolo tipas, siuntėjo ir gavėjo IP adresai, prievadų numeriai, srauto dydis, trukmė, siunčiamų ir gaunamų paketų ilgis, segmento dydis ir t. t. Rinkiniai skiriasi stebėtų atakų tipais ir surinkimo būdu. Abiejuose duomenų rinkiniuose dalis rinkinio atributų yra kategorinio tipo, pavyzdžiui, protokolo tipas, paketų tipai yra kategorijos tipo, todėl prieš naudojant tokie duomenys buvo transformuoti į skaitines reikšmes. Atliktas duomenų valymas, ištrinant nevisiškai užpildytus įrašus, taip pat duomenys normalizuoti naudojant „Min-Max“ algoritmą. Tyrime duomenų rinkiniai buvo padalinti į duomenis, skirtus mokymui 70 % ir testavimui 30 %. Atliekant kryžminį validavimą, šis santykis kito nuo 10 iki 90 %. Pagrindinės duomenų rinkinių charakteristikos pateiktos 1 lentelėje.

1 lentelė. Duomenų rinkinių pagrindinės charakteristikos
Table 1. Main characteristics of datasets

Charakteristikos	CSE-CIC-IDS 2018	UNSW-NB 15
Atakų scenarijų kiekis	7	9
Atributų skaičius	80	49
Klasių skaičius	15	10
Įrašų skaičius	16 232 945	257 673
Dydis, GB	6,41	100
Naudoti įrankiai	Pcap	Pcap, tcpdump
Žymėjimas	Taip	Taip
Balansavimas	Ne	Ne

1.2. Modelių sudarymas

Siekiant ištirti mašininio mokymosi metodų tikslumą ir turėti galimybę palyginti gautus rezultatus su pasiūlytais hibridiniais metodais, buvo pasirinkti 6 mašininio mokymo algoritmai, t. y. sprendimų medžio (angl. *Decison Tree*), atsitiktinio miško (angl. *Random Forest*), naivaus Bajeso (angl. *Naive Bayes*), atraminių vektorių mašinos (angl. *Support Vector Machine*), tikimybinio neuroninio tinklo (angl. *Probabilistic Neural Network*) ir daugiasluoksnio perceptrono (angl. *Multilayer Perceptron*). Tokį metodų pasirinkimą lėmė atlikta literatūros apžvalga ir kitų autorių gauti tyrimų rezultatai.

Šiame darbe buvo pasiūlyti ir trys nauji hibridiniai mašininio mokymosi metodai, sudaryti kaip rinkiniai iš trijų anksčiau išvardintų algoritmų. Algoritmai buvo parinkti atsižvelgiant į gautus atskirų metodų tyrimo rezultatus. Sudaryti tokie hibridiniai metodai: „Decision Tree“, „Multi Layer Perceptron“, „Support Vector Machine“ (HM1), „Decision Tree“, „Naive Bayes“, „Multi Layer Perceptron“ (HM2), „Decision Tree“, „Naive Bayes“ ir „Support Vector Machine“ (HM3).

Tyrimui buvo naudojamas KNIME *Analytics* programinis paketas ir kompiuteris su Intel(R) Core(TM) i7-8665U procesoriumi, 32 GB operatyviąja atmintimi, 1TB SSD disku, *Windows 10 Professional* operacine sistema. Atlikus modelio apmokymą buvo analizuojami rezultatai ir tikslinamos parametrų reikšmės.

2. Eksperimentinis tyrimas ir rezultatai

Šiame skyriuje pateikiami mašininio mokymosi algoritmų testavimo rezultatai, nustatomi svarbiausi duomenų rinkinių atributai, pasiūlomi ir tiriami hibridiniai mašininio mokymosi metodai. Gauti rezultatai apibendrinami ir padaromos atitinkamos išvados.

2.1. Mašininio mokymosi algoritmų tyrimas

Pirmajame tyrimo etape buvo atlikti eksperimentai taikant 6 ankstesniame skyriuje minėtus mašininio mokymosi metodus tiriant jų tikslumą ir apmokymo laiką. Tyrime buvo naudojami abu anksčiau aprašyti duomenų rinkiniai. Buvo naudojamos standartinės algoritmų hiperparametrų reikšmės, nustatytos KMINE programiniame pakete (Fillbrunn et al., 2017). Tikslumo skaičiavimo rezultatai pateikti 2 lentelėje. Iš pateiktų skaičiavimų matyti, kad geriausius rezultatus pasiekia daugiasluoksnis perceptrono (MLP) algoritmas. Vidutinis tikslumas siekia 96,42 %. Toliau eina atsitiktinio medžio (RF) ir tikimybinis neuroninio tinklo algoritmai su tikslumu 94,37 % ir 94,29 %.

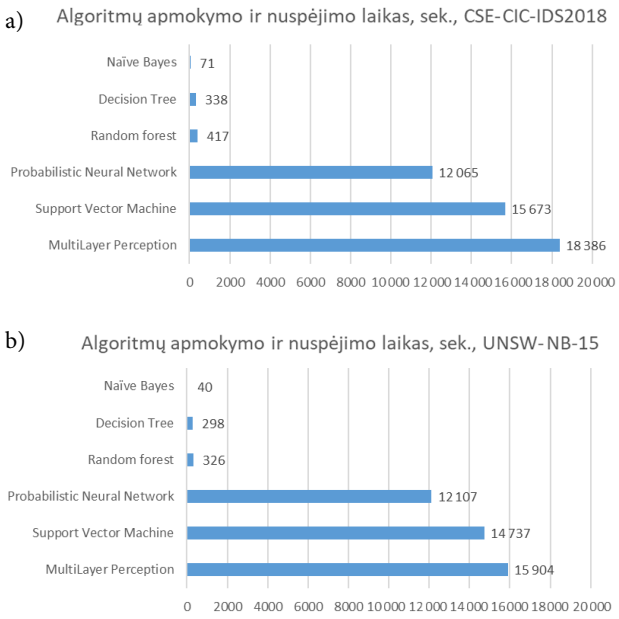
Atlikus analogiškus skaitinių tyrimų testus, naudojant UNSW-NB-15 duomenų rinkinį, buvo gauti žymiai prastesni klasifikavimo tikslumo įverčiai (3 lentelė). Aukščiausias gautas vidutinis tikslumas siekė 79,45 % ir buvo gautas naudojant atsitiktinio medžio algoritmą. Toliau išsidėstė daugiasluoksnio perceptrono (74,34 %) ir sprendimų medžio (71,43 %) algoritmai. Šie tyrimai parodė, kad klasifikavimo tikslumas priklauso nuo duomenų rinkinio ir nuo mašininio mokymosi algoritmo.

2 lentelė. Klasifikavimo tikslumo įverčiai, kai naudojamas CSE-CIC-IDS2018 duomenų rinkinys
Table 2. Classification accuracy obtained using the CSE-CIC-IDS2018 dataset

Nr.	Klasės pavadinimas	RF	DT	SVM	PNN	MLP	NB
1.	Bot	0,9784	0,9992	0,7485	0,9874	0,9889	0,4567
2.	Benign	0,9829	0,9828	0,9710	0,9859	0,9842	0,7499
3.	DoS attacks-SlowHTTPTest	0,8249	0,8228	0,6661	0,7570	0,7230	0,2500
4.	DoS attacks-Hulk	0,9994	0,9934	0,9887	0,9998	0,9987	0,1529
5.	Brute Force –Web	0,8270	0,8447	0,9359	0,8942	0,9973	1,0000
6.	Brute Force –XSS	0,9915	0,8735	0,3590	0,9672	0,9912	0,0100
7.	SQL Injection	0,9211	0,7317	0,4038	0,9583	0,9924	0,1344
8.	DDoS attacks-LOIC-HTTP	0,9995	0,9975	0,8299	0,9991	0,8082	0
9.	Infilteration	0,8866	0,3122	0,7858	0,9261	0,9932	0,9940
10.	DoS attacks-GoldenEye	0,9864	0,9655	0,5272	0,9896	0,9989	0,2043
11.	DoS attacks-Slowloris	0,9995	0,9971	0,9747	0,9992	0,9867	0,4810
12.	FTP-BruteForce	0,7817	0,8129	0,6513	0,6943	1,0000	0,0180
13.	SSH-Bruteforce	0,9955	0,9992	0,9065	1,0000	1,0000	0,0211
14.	DDOS attack-HOIC	0,9973	0,9991	0,9340	0,9986	1,0000	0
15.	DDOS attack-LOIC-UDP	0,9845	0,9869	0,9867	0,9861	1,0000	0,0816
Vidurkis		0,9437	0,8879	0,7779	0,9429	**0,9642**	0,3503

3 lentelė. Klasifikavimo tikslumo įverčiai, kai naudojamas UNSW-NB-15 duomenų rinkinys
Table 3. Classification accuracy obtained using the UNSW-NB-15 dataset

Nr.	Klasės pavadinimas	RF	DT	SVM	PNN	MLP	NB
1.	Normal	0,9338	1,0000	0,9922	0,8514	0,9981	1,0000
2.	Reconnaissance	0,9793	0,9038	0,3744	0,7708	0,9181	0
3.	Backdoor	0,4002	0,2291	0,0466	0,2047	0,9395	0,0319
4.	DoS	0,9691	0,5393	0,4800	0,8286	0,6168	0,1545
5.	Exploits	0,7685	0,7265	0,8003	0,8432	0,5935	0,7555
6.	Analysis	0,7960	0,8250	0,0000	0,2400	0,0000	0,0099
7.	Fuzzers	0,9121	0,8595	0,4806	0,5652	0,4054	0,2473
8.	Worms	0,7667	0,6842	0,0123	0,0000	0,9997	0,0017
9.	Shellcode	0,4187	0,3835	0,0539	0,0944	0,9714	0,0124
10.	Generic	1,0000	0,9921	0,9137	0,9647	0,9913	0,8523
Vidurkis		**0,7945**	0,7143	0,4154	0,5363	0,7434	0,3066



2 paveikslas. Mašininio mokymosi modelių mokymosi laikas sekundėmis, kai naudojami UNSW-NB-15 (a) ir UNSW-NB-15 (b) duomenų rinkiniai
Figure 2. Training time in seconds for machine learning models using UNSW-NB-15 (a) and UNSW-NB-15 (b) datasets

Tyrimų metu buvo skaičiuojamas ir skirtingų modelių apmokymo laikas, nes tai yra svarbus kriterijus, kalbant apie modelių pakartotinį apmokymą realiose įsilaužimų aptikimo sistemose (2 pav.). Analizuojant metodų apmokymo laikus, matyti, kad greičiausiai apsimoko naivus Bajeso modelis, toliau eina sprendimų medžio ir atsitiktinio miško algortimai. Ilgiausiai apsimoko daugiasluoksnis perceptronas. Skirtumas tarp greičiausio ir lėčiausio apmokymo siekia 397 karto. Matome, kad apmokymo laikai mažai priklauso nuo duomenų rinkinio, t. y. tos pačios apmokymo laikų tendencijos išlieka abiejuose duomenų rinkiniuose.

2.2. Svarbiausieji duomenų rinkinio atributai

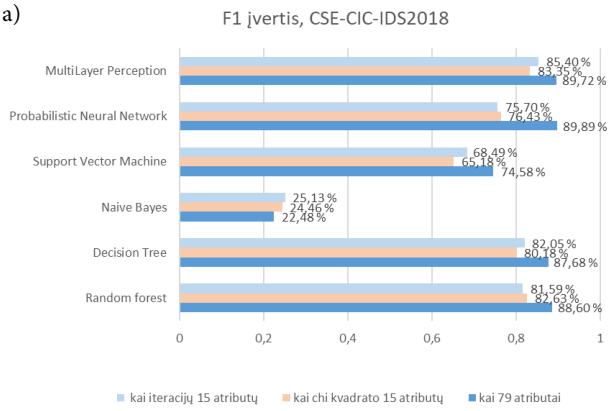
2.1 skyriuje atlikti eksperimentiniai mašininio mokymo metodų tyrimai buvo vykdomi naudojant visus duomenų rinkinių atributus. Tačiau kitų autorių atliktuose moksliniuose darbuose pažymima, kad galima sumažinti duomenų rinkinio atributų skaičių, pašalinant nesvarbius kintamuosius, ir kartu pasiekti didesnį modelio našumą, neprarandant metodo tikslumo (Chen et al., 2020). Šiam tikslui pasiekti taikomas Chi kvadrato χ^2 testas, kuris naudojamas statistikoje norint patikrinti dviejų įvykių nepriklausomumą, ir iteracijų metodas. Mašininio mokymosi procese taikant Chi kvadrato χ^2 testą galima nustatyti, kaip modelis lygina stebimus ir realius duomenis, atributų priklausomumą, ir atrinkti geriausius atributus modeliui sudaryti. Kuo atributai daugiau priklausomi, tuo Chi kvadrato χ^2 reikšmė didesnė. Tiriant modelį, kai naudojami tik atrinkti svarbiausi atributai, galima tikėtis ir didesnio klasifikavimo tikslumo (Pandis, 2016). Svarbiausių atributų atranka atlikta su CSE-CIC-IDS2018 ir UNSW-NB-15 duomenų rinkiniais ir šešiais mašininio mokymo algoritmais. Penkiolika svarbiausių atributų iš kiekvieno duomenų rinkinio pateikta 4 lentelėje. Buvo atlikti analogiški svarbiausių atributų nustatymo tyrimai taikant iteracinį algoritmą. Daugiau nei 50 % gautų svarbiausių atributų buvo tokie patys, tačiau jų reitingo vieta sąraše skiriasi.

4 lentelė. Taikant Chi kvadratu χ^2 testą gauti svarbiausi atributai, pateikti mažėjimo tvarka
Table 4. The principal attributes obtained by applying the Chi-squared $\chi 2$ test and presented in descending order

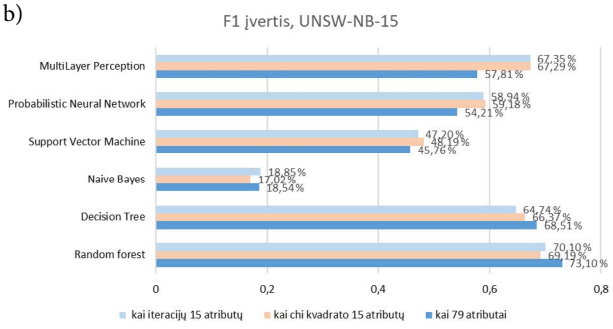
Nr.	Atributo pavadinimas	
	CSE-CIC-IDS2018	UNSW-NB-15
1.	Flow IAT Max	sload
2.	Flow Duration	sbytes
3.	Fwd Pkt Len Max	smean
4.	Fwd Pkts/s	dmean
5.	Flow IAT Mean	dbytes
6.	Fwd IAT Std	dur
7.	Fwd Seg Size Avg	ct_srv_src
8.	Fwd Pkt Len Mean	ct_srv_dst
9.	Pkt Size Avg	dpkts
10.	TotLen Fwd Pkts	ct_dst_sport_ltm
11.	Active Min	rate
12.	Fwd Seg Size Avg	ct_dst_src_ltm
13.	Active Mean	service
14.	Pkt Len Max	ct_dst_ltm
15.	Bwd Seg Size Avg	ct_state_ttl

Siekiant nustatyti, kaip atrinkti svarbiausi atributai veikia mašininio mokymo algoritmų tikslumą, buvo atlikti šešių mašininio mokymo algoritmų bandymai su duomenų rinkiniais CSE-CIC-IDS2018 ir UNSW-NB-15. Algoritmų modeliams buvo naudojamas sumažintas 15 atributų duomenų rinkinys, o gauti rezultatai palyginti naudojant F1 metrikas, gautas atliekant skaičiavimus su pilnais duomenų rinkiniais (3 pav.). Analizuojant rodiklio F1 reikšmes matyti, kad naudojant sumažintus duomenų rinkinius F1 reikšmė sumažėjo nuo 4 % iki 16 %, kai F1 reikšmė buvo didesnė nei 50 %. Esant mažesnėms nei 50 % F1 reikšmėms, tikslumą pavyko padidinti. Taip pat buvo matuojamas algoritmų apmokymo laikas, naudojant rinkinius su mažesniu duomenų atributų skaičiumi. Apmokymo laikas sumažėjo visiems algoritmams nuo 4 % iki 48 %. Analizuojant rezultatus, matyti, kad F1 reikšmė skiriasi 2–3 %, kai naudojami duomenų atributai, atrinkti Chi kvadrato testu, ir atributai, gauti iteracijų metodu. Mažą skirtumą veikia tų pačių atributų pasikartojimas abiejuose rinkiniuose. Taip pat galima teigti, kad iteracijų metodas leidžia pasiekti gerus rezultatus kaip ir Chi kvadrato χ^2 testas.

a)



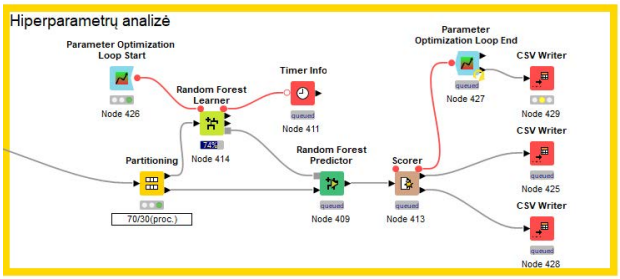
b)



3 paveikslas. Rezultatų su svarbiausiais atributais palyginimas, kai naudojamas duomenų rinkinys CSE-CIC-IDS2018 (a) ir UNSW-NB-15 (b)
Figure 3. Comparison of the results using the principal attributes of the dataset CSE-CIC-IDS2018 (a) and UNSW-NB-15 (b)

2.3. Hiperparametrų tyrimas

Mašininio mokymosi algoritmai parametrizuojami hiperparametrais, kurie veikia klasifikatorių tikslumą. Keičiant hiperparametrų reikšmes, galima padidinti modelio tikslumą, greitį ir gauti geresnius rezultatus. Algoritmų hiperparametrų optimizavimas KNIME *Analytics* aplinkoje



4 paveikslas. Hipeparametrų optimizavimo komponentų sujungimo schema
Figure 4. Component diagram used for optimization of hypeparameters

atliekamas pagal 4 pav. pateiktą komponentų sujungimo schemą. KNIME *Analytics* aplinkoje sukurti algoritmų hiperparametrų optimizavimo modeliai ir atlikti eksperimentai su kiekvienu algoritmu parenkant hiperparametrų reikšmių iteracijų skaičių ir žingsnį. Toliau sukurti parametrų kintamieji, kuriems buvo nusatyti hiperparametrų reikšmių intervalai ir žingsnis. Kiekvienos iteracijos metu algoritmai buvo apmokomi ir testuojami, o rezultatai kaupiami CSV failuose.

Hiperparametrų reikšmėms rasti buvo apibrėžti iteracijų intervalai ir žingsniai. Pasirinktos pradinės reikšmės, nuo kurių pradedama kiekviena nauja iteracija. Pasibaigus bandymui su vienu parametru, grąžinamos pradinės reikšmės, nustatoma naujo hiperparametro reikšmė, su kuriuo atliekami tolesni bandymai. Optimalios hiperparametrų reikšmės ir modelių tikslumas, gautas su šiomis reikšmėmis, kai buvo nagrinėtas CSE-CIC-IDS2018 duomenų rinkinys, pateiktas 5 lentelėje. Labai panašūs rezultatai buvo gauti ir dirbant su antruoju rinkiniu. Analizuojant rezultatus matyti, kad pavyko pagerinti atsitiktinio miško algoritmo tikslumą 25 %, visais kitais atvejais tikslumas pagerėjo iki 1 %. Pakeitus hiperparametrų reikšmes, padidėjo sprendimų medžio ir atsitiktinio miško modelio apmokymo laikas. Kitų algoritmų apmokymo laikai liko nepakitę.

5 lentelė. Eksperimentų su hiperparametrais rezultatai duomenų rinkiniui CSE-CIC-IDS2018
Table 5. Investigation results of hyperparameters for the CSE-CIC-IDS2018 dataset

Algoritmas	Hiperparametras	Hiperparametro reikšmė	Geriausias tikslumas	Pradinis tikslumas	Geriausios reikšmės laikas, s	Pirmos reikšmės laikas, s
RF	Medžio gylis	18	0,98425	0,72362	422	197
RF	Modelių skaičius	176	0,98242	0,98179	574	427
DT	Įrašų skaičius mazge	7	0,97426	0,96931	335	325
DT	Gijų skaičius	2	0,97251	0,97273	327	327
SVM	*Sigma* parametras	1	0,94150	0,94150	15572	15572
PNN	*Theta* plus parametras	0,4	0,99031	0,99031	12141	12141
PNN	*Theta* minus parametras	0,2	0,98123	0,98120	12366	12366
MLP	Paslėptų sluoksnių skaičius	1	0,98022	0,98022	18536	18536
MLP	Paslėptų neuronų skaičius sluoksnyje	12	0,97681	0,97493	20503	18096

2.4. Hibridiniai metodai

Siekiant pagerinti įsilaužimų aptikimo sistemų tikslumą, buvo pasiūlyti trys nauji hibridiniai mašininio mokymosi metodai (žr. 1.2 skyrių). Pasiūlytų metodų efektyvumas buvo tiriamas naudojant tuos pačius du pilnus duomenų rinkinius. Buvo skaičiuojamas klasifikavimo tikslumas. Rezultatai pateikti 5 pav. Hibridinių modelių klasifikavimo tikslumo įverčiai CSE-CIC-IDS2018 duomenų rinkiniui yra didesni nei bandymuose su UNSW-NB-15 duomenimis. Taip yra dėl duomenų rinkinių skirtumų, atributų turimos naudingos informacijos, taip pat dėl to, kad hibridinius modelius sudarančių algoritmų rezultatų įverčiai ankstesniuose bandymuose su duomenų rinkiniu CSE-CIC-IDS2018 buvo didesni. Gauti rezultatai rodo, kad suformuotų hibridinių modelių taikymas yra naudingas siekiant padidinti modelio tikslumą ir našumo klasifikuojant anomalijas. Didžiausias tikslumas 99,34 % buvo pasiektas taikant HM2 metodą. Tai 3,13 % tiksliau nei geriausią tikslumą parodęs tikimybinis neuroninis tinklas, testuotas panaudojant optimizuotas hiperparametrų reikšmes. Mažiausia modelio apmokymo trukmė gauta taikant HM3 hibridinį metodą.

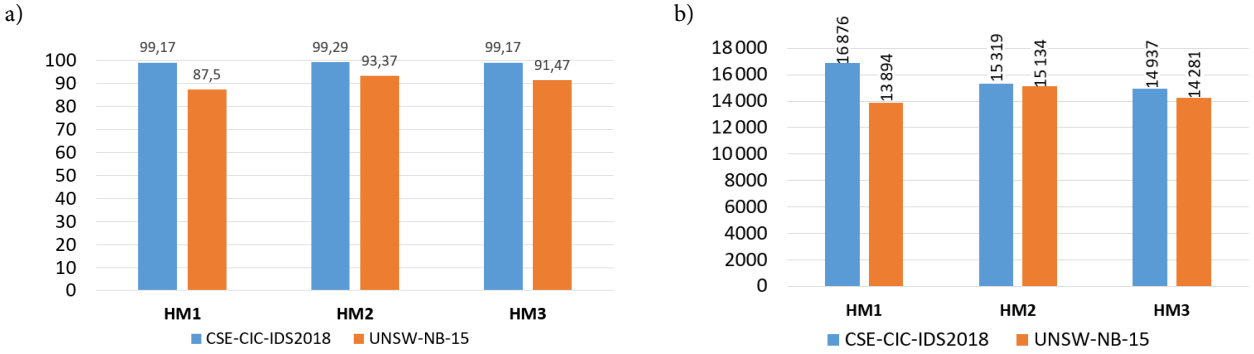
2.5. Metodų reitingavimas

Darbe pasiūlytų hibridinio mokymosi metodų ir nagrinėtų mašininio mokymosi algoritmų palyginimui bei įvertinimui atlikti buvo pasirinktos trys reitingavimo metodikos (Vaitkevicius & Marcinkevicius, 2020), t. y. standartinis reitingas (angl. *Standard competition rating*), tankusis reitingas (angl. *Dense rating*) ir dalinis reitingas (angl. *Fractional rating*). Reitinguojant pagal standartinį reitingą, vienodi matavimai gauna tą patį reitingą, bet kitas reitingo numeris prasideda praleidus tiek skaičių, kiek buvo prieš tai vienodą rangą gavusių matavimų. Tankus reitingas veikia panašiai kaip standartinis, bet reitingų numeriai eina iš eilės be jokių trūkių. Dalinis reitingas naudoja aritmetinį vidurkį vienodiems matavimams reitinguoti. Prieš taikant vieną iš reitingavimo metodikų, rezultatai surūšiuojami pagal atitinkamą vertinimo kriterijų, tuomet pritaikoma atitinkama metodika. Kiekvienai reitingavimo metodikai paskirstomi taškai nuo aukščiausio įvertinimo 10 iki žemiausio 1. Taškai skaičiuojami pagal tokią formulę:

$$Taškai_i = N_m - Reitingas_i + 1, \quad (1)$$

čia N_m – nagrinėtų mašininio mokymosi metodų skaičius; $Reitingas_i$ – metodo reitingas.

Darbe nagrinėtiems mašininio mokymosi metodams palyginti ir reitinguoti buvo pasirinktos tokios metrikos – tikslumas, preciziškumas, F1 kriterijus ir modelio apmokymo laikas. 6 lentelėje pateikti visų nagrinėtų metodų tikslumo reitingai ir pagal tris metodikas suskaičiuoti reitingo taškai. Matome, kad visi trys pasiūlyti hibridiniai metodai leidžia pasiekti aukščiausią tikslumą.



5 paveikslas. Hibridinių metodų tikslumo (a) ir apmokymo laiko (b) palyginimas
Figure 5. Comparison of accuracy (a) and training time (b) of hybrid methods

6 lentelė. Tikslumo reitingai CSE-CIC-IDS2018 duomenų rinkiniui
Table 6. Accuracy rankings for the CSE-CIC-IDS2018 dataset

Tikslumas	SCR	FR	DR	Algoritmas	SCR taškai	FR taškai	DR taškai
0,9939	1	2	1	HM 2	9	8	9
0,9931	1	2	2	HM 1	9	8	8
0,9791	1	2	2	HM 3	9	8	8
0,9632	4	4,5	3	Sprendimų medis	6	5,5	7
0,9548	4	4,5	3	Atsitiktinis miškas	6	5,5	7
0,9219	6	5	4	Atraminių vektorių mašina	4	5	6
0,8928	7	7,5	5	Tikimybinis neuroninis tinklas	3	2,5	5
0,8855	7	7,5	5	Daugiasluoksnis perceptronas	3	2,5	5
0,3449	9	9	6	Naivus Bajeso	1	1	4

Analogiški reitingo taškai buvo suskaičiuoti preciziškumui, F1 kriterijui ir modelio apmokymo laikui. Šie reitingai buvo sudaryti abiem duomenų rinkiniams. Galiausiai suskaičiuoti suminiai reitingo taškai. Rezultatai pateikti 7 ir 8 lentelėse. Daugiausiai reitingo taškų, tiriant CSE-CIC-IDS2018 duomenų rinkinį, surinko sprendimų medžio algoritmas, kurio rezultatai pagal visas metrikas buvo pakankamai aukšti. Trečia vieta atiteko hibridiniams modeliams HM1 ir HM2. Analizuojant duomenų rinkinį UNSW-NB-15 geriausias buvo atsitiktinio miško algoritmas, toliau sprendimų medžio ir HM1 hibridinis metodas (8 lentelė).

7 lentelė. Jungtinis reitingas CSE-CIC-IDS2018 duomenų rinkiniui
Table 7. Combined ranking for the CSE-CIC-IDS2018 dataset

Algoritmas	SCR taškai	FR taškai	DR taškai
Sprendimų medis	39	37	41
Tikimybinis neuroninis tinklas	35	33,5	37
Atsitiktinis miškas	34	31,5	37
HM1	30	28	32
HM 2	26	24	32
Daugiasluoksnis perceptronas	25	22,5	29
HM3	21	19,5	28
Atraminių vektorių mašina	17	17	26
Naivus Bajeso	13	13	25

8 lentelė. Jungtinis reitingas UNSW-NB-15 duomenų rinkiniui
Table 8. Combined ranking for the UNSW-NB-15 dataset

Algoritmas	SCR taškai	FR taškai	DR taškai
Atsitiktinis miškas	43	42,5	43
Sprendimų medis	33	31	35
HM 1	30	28,5	33
HM 2	32	31	32
HM 3	27	25	29
Daugiasluoksnis perceptronas	26	24,5	28
Tikimybinis neuroninis tinklas	18	16,5	24
Naivus Bajeso	15	14,5	23
Atraminių vektorių mašina	12	11,5	20

Išvados

1. Atlikti eksperimentiniai tyrimai analizuojant ir tiriant tris pasiūlytus hibridinio mokymosi metodus ir papildomai šešis kitus mašininio mokymosi algoritmus panaudojant viešai prieinamus duomenų rinkinius CSE-CIC-IDS2018 ir UNSW-NB-15. Nustatyta, kad algoritmų tikslumas ir modelių apmokymo laikas priklauso nuo duomenų rinkinio kokybės, dydžio ir atributų skaičiaus.
2. Svarbiausių atributų atranka duomenų rinkiniuose sumažina anomalijų aptikimo tikslumą, kai jis yra didesnis nei 50 %. Kartu sumažinamas ir mašininio mokymosi modelių apmokymo laikas. Svarbiausių atributų atrankos rezultatai priklauso nuo duomenų rinkinyje esančių atributų skaičiaus ir duomenų. Kai rinkinyje daug triukšmingų duomenų, svarbių atributų atranka padidina klasifikavimo greitį.
3. Atlikus algoritmų hiperparametrų tobulinimą, nustatyta, kad 25 % buvo pagerintas atsitiktinio miško algoritmo tikslumas. Visais kitais atvejais tikslumas pagerėjo iki 1 %. Pakeitus hiperparametrų reikšmes padidėjo tik sprendimų medžio ir atsitiktinio miško modelio apmokymo laikas, kitų algoritmų apmokymo laikai liko nepakitę.
4. Suformuotų trijų hibridinių modelių bandymų rezultatai rodo, kad hibridiniai modeliai leidžia padidinti įsilaužimų tikslumą. Didžiausias tikslumas 99,34 % buvo pasiektas taikant HM2 metodą. Tai 3,13 % tiksliau nei geriausią tikslumą parodęs tikimybinis neuroninis tinklas, testuotas panaudojant optimizuotas hiperparametrų reikšmes.

Parama

Šis tyrimas negavo jokios paramos ar finansavimo.

Autorių indėlis

Metodologija K. Č., D. S.; metodų analizė K. Č., D. S.; tyrimas K. Č., D. S.; straipsnio rašymas K. Č.; rezultatų vizualizavimas K. Č., D. S.; projekto administravimas K. Č.

Literatūra

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, *60*, 19–31. https://doi.org/10.1016/j.jnca.2015.11.016

Aminanto, E., & Kim, K. (2016). Deep learning in intrusion detection system: An overview. In 2016 *International Research Conference on Engineering and Technology* (*2016 IRCET*) [Conference presentation]. Higher Education Forum.

Ashiku, L., & Dagli, C. (2021). Network intrusion detection system using deep learning. *Procedia Computer Science*, *185*, 239–247. https://doi.org/10.1016/j.procs.2021.05.025

Atefinia, R., & Ahmadi, M. (2021). Network intrusion detection using multi-architectural modular deep neural network. *The Journal of Supercomputing*, *77*(4), 3571–3593. https://doi.org/10.1007/s11227-020-03410-y

Bao, Y., Tang, Z., Li, H., & Zhang, Y. (2019). Computer vision and deep learning–based data anomaly detection method for structural health monitoring. *Structural Health Monitoring*, *18*(2), 401–421. https://doi.org/10.1177/1475921718757405

Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2014). Towards an unsupervised method for network anomaly detection in large datasets. *Computing and Informatics*, *33*(1), 1–34.

Bulavas, V., Marcinkevičius, V., & Rumiński, J. (2021). Study of multi-class classification algorithms' performance on highly imbalanced network intrusion datasets. *Informatica*, *32*(3), 441–475. https://doi.org/10.15388/21-INFOR457

Chen, R. C., Dewi, C., Huang, S. W., & Caraka, R. E. (2020). Selecting critical features for data classification based on

machine learning methods. *Journal of Big Data*, *7*(1), 1–26. https://doi.org/10.1186/s40537-020-00327-4

Chkirbene, Z., Eltanbouly, S., Bashendy, M., AlNaimi, N., & Erbad, A. (2020). Hybrid machine learning for network anomaly intrusion detection. In *2020 IEEE International Conference on Informatics*, *IoT*, *and Enabling Technologies* (*ICIoT*) (pp. 163–170), Doha, Qatar. https://doi.org/10.1109/ICIoT48696.2020.9089575

Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, *50*, 102419. https://doi.org/10.1016/j.jisa.2019.102419

Fillbrunn, A., Dietz, C., Pfeuffer, J., Rahn, R., Landrum, G. A., & Berthold, M. R. (2017). KNIME for reproducible cross-domain analysis of life science data. *Journal of Biotechnology*, *261*, 149–156. https://doi.org/10.1016/j.jbiotec.2017.07.028

Hassan, M. M., Gumaei, A., Alsanad, A., Alrubaian, M., & Fortino, G. (2020). A hybrid deep learning model for efficient intrusion detection in big data environment. *Information Sciences*, *513*, 386–396. https://doi.org/10.1016/j.ins.2019.10.069

Kanimozhi, V., & Jacob, T. P. (2019). Artificial intelligence based network intrusion detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. In 2019 *International Conference on Communication and Signal Processing* (*ICCSP*) (pp. 0033–0036). IEEE. https://doi.org/10.1109/ICCSP.2019.8698029

Kwon, D., Kim, H., Kim, J., Suh, S. C., Kim, I., & Kim, K. J. (2019). A survey of deep learning-based network anomaly detection. *Cluster Computing*, *22*(1), 949–961. https://doi.org/10.1007/s10586-017-1117-8

Megantara, A. A., & Ahmad, T. (2021). A hybrid machine learning method for increasing the performance of network intrusion detection systems. *Journal of Big Data*, *8*(1), 1–19. https://doi.org/10.1186/s40537-021-00531-w

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of machine learning*. The MIT Press.

Pandis, N. (2016). The chi-square test. *American Journal of Orthodontics and Dentofacial Orthopedics*, *150*(5), 898–899. https://doi.org/10.1016/j.ajodo.2016.08.009

Trustware. (2020). *Trustware global security report*. https://www.trustwave.com/en-us/resources/library/documents/2020-trustwave-global-security-report/

Vaitkevicius, P., & Marcinkevicius, V. (2020). Comparison of classification algorithms for detection of phishing websites. *Informatica*, *31*(1), 143–160. https://doi.org/10.15388/20-INFOR404

NETWORK INTRUSION DETECTION USING HYBRID MACHINE LEARNING METHODS

K. Čiurlienė, D. Stankevičius

Abstract

Network intrusion detection is a relevant cybersecurity research field. The growing number of intrusions requires more sophisticated methods to protect computer networks. Various machine learning algorithms are used to detect network intrusions and anomalies, but their accuracy is limited. In this research, we address the problem of improving network-level intrusion detection by applying hybrid machine-learning algorithms. The paper proposes three new hybrid machine learning methods and investigates their accuracy using two publicly available datasets CSE-CIC-IDS2018 and NSW-NB-15. In order to increase the accuracy of the classification models, hyperparameter optimization was performed. The iteration method and the Chi-square $\chi 2$ test were used to identify significant features of the data set. Analyzing the research results, it was found that the highest network anomaly recognition accuracy of 99.34% was achieved by applying a hybrid algorithm consisting of a decision tree, naive Bayesian, and multilayer perceptron algorithms. Achieved result is 3.13% higher than the best accuracy achieved by individual machine learning algorithms. In order to comprehensively evaluate the studied machine learning algorithms and their suitability for detecting intrusions in a computer network, the algorithms were ranked using the SCR, DR, FR ranking methods.

Keywords: network anomalies, machine learning, $\chi 2$- Chi-squared test, hyperparameters, hybrid algorithms.