Vilniaus Universitetas
Duomenų mokslo ir skaitmeninių technologijų institutas
Kognityvinių skaičiavimų grupė

Ataskaitinė informatikos krypties doktorantų konferencija

Veiklos ataskaita už 2021-03 – 2021-09

Žydrūnas Vaišnoras (DMSTI-DS-N009-20-11)

2021-09-30, Vilnius, Lietuva

Doktorantūros studijos

- Disertacijos pavadinimas – Mašininio mokymosi metodų vystymas įsilaužimams aptikti kompiuterių tinkluose
- Doktorantas – Žydrūnas Vaišnoras (DMSTI-DS-N009-20-11)
- Darbo vadovė – prof. dr. Olga Kurasova
- Mokslo kryptis – 09 P Informatika
- Doktorantūros laikotarpis – 2019-2023 m.
- Studijų metai – 2021 m. (II)

Tyrimo uždaviniai

- Atlikti skirtingų mašininio mokymosi metodų, naudojamų kompiuterių tinklo anomalijoms atpažinti, analizę ir tyrimą;
- Parinkti tyrimo metodiką iškeltiems uždaviniams spręsti;
- Sukurti našesnį mašininio mokymosi metodą anomalijoms atpažinti realaus laiko duomenims;
- Pritaikyti sukurtą mašininio mokymosi modelį realaus laiko duomenims ir atlikti gautų duomenų analizę, rezultatų apibendrinimą, išvadų parengimą.

Planuojamas mokslinis naujumas

- Sukurtas našesnis mašininio mokymosi modelis įsilaužimams atpažinti realaus laiko duomenims;
- Sukurtas metodas, kuris naudos kuo įmanoma mažiau „nematytų“ kompiuterių tinklo duomenų paketų mašininio mokymosi modelio ap(si)mokymui – anomalijų atpažinimui;
- Mašininio mokymosi modelis bus pritaikomas darbui virtualioje aplinkoje, konteinerizavimo platformose.

VISŲ STUDIJŲ PLANAS IR JO VYKDYMO SUVESTINĖ

Studijų metai	Egzaminai		Dalyvavimas konferencijose		Publikacijos		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė
I (2019/2020)	1	1					
II (2020/2021)	**2**	**2**	**1**	0			
III (2021/2022)	1		1 + 1 (skola ir II metų)		1		
IV (2022/2023)					1		
Iš viso:	4	3	2		2		

ATASKAITINIŲ METŲ DARBO PLANAS IR JO ĮVYKDYMAS

Egzaminai		Dalyvavimas konferencijose		Publikacijos	
Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta
Fundamentalieji informatikos ir informatikos inžinerijos metodai	Išlaikyta: Fundamentalieji informatikos ir informatikos inžinerijos metodai.	Suplanuota konferencija atšaukta dėl COVID-19 pandemijos. Vietoje jos suplanuota dalyvauti ITMS'2021 konferencijoje 2021 m. spalio mėn.	Nesudalyvauta konferencijoje ITMS'2021.		
Gilieji neuroniniai tinklai	Išlaikyta: Gilieji neuroniniai tinklai.		Dalyvauta tarptautinėje doktorantų vasaros/žiemos mokykloje „4th International School on Deep Learning"		

MOKSLINIŲ TYRIMŲ IR DISERTACIJOS RENGIMO ETAPAI (1)

	Darbo pavadinimas	Atlikimo terminai	Pastabos
1.	**Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):** 1.1. Disertacijos tyrimo objekto detalizavimas. 1.2. Atlikti mašininio mokymosi metodų taikymo kompiuterių tinkluose analitinę apžvalgą. 1.3. Nustatyti (identifikuoti) mokslines problemas, kylančias uždaviniuose, susijusiuose su anomalijų aptikimu kompiuterių tinkluose taikant mašininio mokymosi metodus. 1.4. Tyrimo tikslo suformavimas.	2019 m. spalio mėn. – 2020 m. rugsėjo mėn.	Atliktas disertacijos tyrimo objekto detalizavimas, nustatytos mokslinės problemos ir suformuotas tyrimo tikslas.

MOKSLINIŲ TYRIMŲ IR DISERTACIJOS RENGIMO ETAPAI (2)

	Darbo pavadinimas	Atlikimo terminai	Pastabos
2.	Mokslinio tyrimo vykdymas: ---------- 2.1. Tyrimo metodikos sudarymas: 2.1.1. Tyrimo metodikos iškeltiems uždaviniams spręsti parinkimas; 2.1.2. Teorinio ir empirinio tyrimų suplanavimas pagal pasirinktą metodiką. ---------- **2.2. Teorinis tyrimas:** **2.2.1. Mašininio mokymosi metodų, naudojamų kompiuterių tinkluose anomalijoms aptikti, tyrimas.** **2.2.2. Anomalijų atpažinimo mašininio mokymosi metodo sukūrimas ir/ar testavimas.** ---------- 2.3. Empirinis tyrimas: 2.3.1. Sudarytų metodų pritaikymas praktinių uždavinių sprendimui. 2.3.2. Gautų duomenų analizė, rezultatų apibendrinimas, išvadų parengimas.	2020 m. spalio mėn. **2020 m. lapkričio mėn. – 2021 m. rugsėjo mėn.** 2021 m. spalio mėn. – 2022 m. gegužės mėn. 2022 m. birželio mėn. – 2022 m. rugsėjo mėn.	**Atliktas teorinis tyrimas – sukurtas naujas metodas dideliems kompiuterių tinklo duomenims rinkti, kaupti apdoroti ir pritaikyti mašininio mokymosi algoritmams.**

Disertacijos tema, tyrimo objektas ir tikslai

Disertacijos tema:

- Mašininio mokymosi metodų vystymas įsilaužimams aptikti kompiuterių tinkluose.

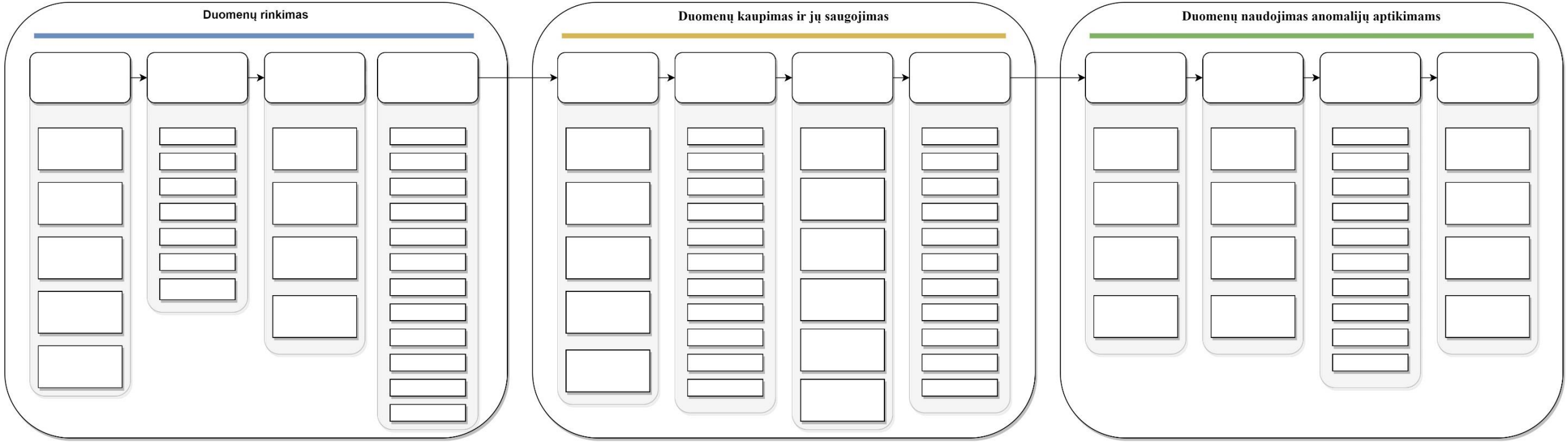
Tyrimo objektas:

- Kompiuterių tinklo įrenginiais sukaupti realaus laiko duomenys;
- Mašininio mokymosi algoritmai įsilaužimams aptikti.

Tyrimo tikslas:

- Išvystyti našesnį mašininio mokymosi algoritmą kibernetiniams įsilaužimams atpažinti kompiuterių tinkluose pritaikant realaus laiko duomenis.

Per pusmetį gautų mokslinių darbų rezultatai (1)



Sisteminės architektūros schema (be teksto)

Sisteminės architektūros schema su tekstu pateikta mokslinėje ataskaitoje. Ji bus viešai paskelbta po išspausdinimo moksliniame straipsnyje.

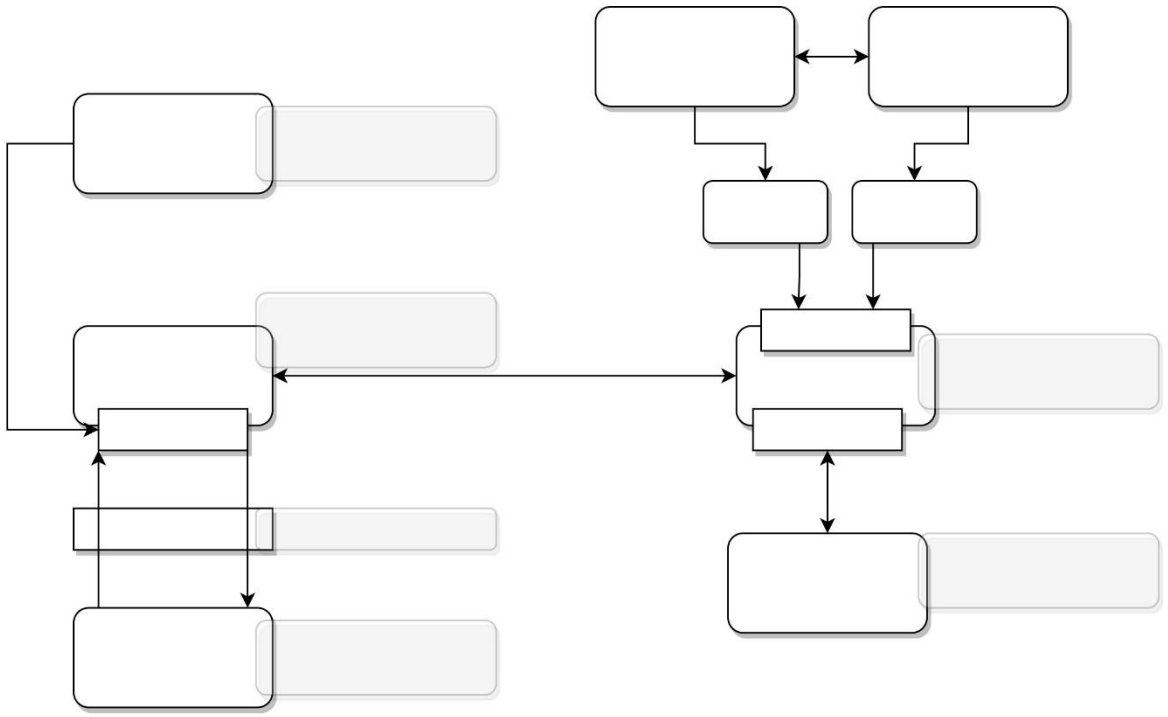
Per pusmetį gautų mokslinių darbų rezultatai (2)

- Analizuota literatūra apie skirtingų mašininio mokymosi metodų taikymą esamiems/sukurtiems kibernetinių įsibrovimų duomenų rinkiniams. Buvo siekiama išanalizuoti/rasti sąryšį tarp kibernetinių atakų aptikimo tikslumo ir duomenų rinkinio (jo sudedamųjų dalių);
- Buvo nagrinėtos kompiuterių tinklo duomenų srauto transliavimo (angl. *stream*) ir apdorojimo varikliai (angl. *proccess engines*) siekiant be uždelsimo, apribojimų ir neprarandant tam tikrą procentą duomenų paketų dėl didžiulių duomenų srauto generavimo duomenų centruose, interneto ryšio tiekėjo pagrindiniuose (angl. *core*) maršrutizatoriuose, komutatoriuose;

Per pusmetį gautų mokslinių darbų rezultatai (3)

- Taip pat buvo nagrinėti didelių duomenų (angl. *big data*) saugojimo ir archyvavimo sistemų, platformų (angl. *framework*) moksliniai straipsniai ir programinės įrangos dokumentacijos siekiant išsiaiškinti kaip greičiau įrašyti, apdoroti, nuskaityti ir suieškoti kaupiamus ir generuojamus kompiuterių tinklo duomenų paketus;
- Išanalizavus mokslinius literatūros šaltinius išsiaiškinta, kad neprogramuojama aparatinė įranga (GPU, TPU, IPU ir kiti) sparčiau atlieka skaičiavimus mašininių mokymosi algoritmuose nei įprasti skaičiavimo kompiuteriai, serveriai. Tai įgalina mašininio mokymosi algoritmus taikant didelius duomenų rinkinius (daug įrašų) apsimokyti greičiau.

Per pusmetį gautų mokslinių darbų rezultatai (4)



Naujojo siūlomo metodo schema (be teksto)

Naujojo siūlomo metodo schema su tekstu pateikta mokslinėje ataskaitoje. Ji bus viešai paskelbta po išspausdinimo moksliniame straipsnyje.

Per pusmetį dalyvauta moksliniuose, bei kitose veiklose

- **Dalyvauta** suplanuotoje Briuselio ir Londono instituto (Institute for Research Development, Training and Advice, IRDTA) dirbtinio intelekto **vasaros stovykloje „4th International School on Deep Learning“**;
- **Dalyvauta** Europos universitetų aljanso (įskaitant Vilniaus universitetą) „Arqus“ (Arqus European University Alliance) organizuotoje kibernetinio saugumo **vasaros stovykloje „2021 Summer School in Cybersecurity“**;
- **Dalyvauta** kasmetiniame Lietuvos kompiuterininkų sąjungos organizuojamoje **konferencijoje „Kompiuterininkų dienos – 2021“**;
- **Dalyvauta** debesų kompiuterijos ir kibernetinio saugumo **konferencijoje „IT CONNECTED“.**

Bendrųjų gebėjimų ugdymas

- Mokslinių rezultatų publikavimas pagal formalaus vertinimo reikalavimus – 0,1 ECTS (2020-11-10);
- Atvirosios prieigos kompetencijų tobulinimas – 0,2 ECTS (2020-11-10);
- Viso: 0,3 ECTS.

Kita pusmečio veikla (neakademiniai)

- Dalyvauta kompiuterių tinklų įrangos gamintojo CISCO konferencijoje „CISCO LIVE! 2021“, kurios metu didžiausias dėmesys buvo skiriamas debesų kompiuterijos saugumo užtikrinimui, vartotojų autentifikacijos ir stabilumu užtikrinimui jungiantis/ryšiams su debesų kompiuterija. Pasiūlyti keli nauji produktai saugiam nuotoliniam darbui dėl pandeminės situacijos aplinkybių.

Mokslinių darbai kitam pusmečiui (1)

- Bus gilinimąsi ir siekiama susisteminti kompiuterių tinklo įprastų ir kenkėjiškų duomenų srautų pagrindinių požymių skirtumus, jų klasifikavimo problematikas;
- Bus analizuojama mokslinė literatūra apie duomenų konteinerizavimo ir mikrosegementavimo procesus, duomenų srauto judėjimą juose ir jų saugumą;
- Bus siekiama išvystyti našesnį mašininio mokymosi algoritmą pritaikant teorinio tyrimo metu sukurtą metodą, kuris gebėtų išsaugoti didžiulius generuojamus srautus, apdoroti/paruošti sukauptus duomenis ir pritaikyti juos mašininio mokymosi algoritmams siekiant nustatyti nematytas (angl. zero day) kibernetines atakas kompiuterių tinkle.

Vilnius University

Ačiū