

Įrodymais grįsto tinklo srauto duomenų modeliavimas ir analizė kibernetinių incidentų užkardymui

Ataskaita už 2025-2026 mokslo metų I pusmetį. Doktorantūros pradžios ir pabaigos
metai: 2023-2027

Doktorantas: Virgilijus Krinickij
Darbo vadovas: Prof. Dr. Linas Bukauskas

Vilniaus universitetas
Matematikos ir informatikos fakultetas
Informatikos institutas
Kibernetinio saugumo laboratorija

2026-03-31

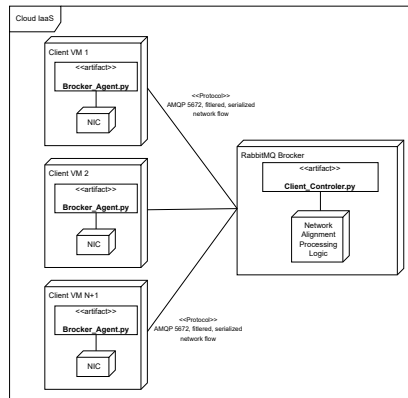
- 1 Išlaikytas pasirenkamasis dalykas „Tinklų modeliai ir algoritmai“.
- 2 Patobulinta aplinka eksperimentams ir tinklo srauto gavybai.
- 3 Modelio pritaikymas eksperimentams pasinaudojant patobulinta aplinka.
- 4 Išsiųstas straipsnis į „IEEE Access“ žurnalą.
- 5 Gautos recenzijos iš „IEEE Access“. Straipsnis tvarkomas pagal recenzijas.
- 6 Sutvarkytas straipsnis išsiųstas į „Baltic Journal of Modern Computing“ žurnalą.

- **Tyrimo objektas** – Duomenų srautų gautu asinchroninio įrašymo metu modeliavimas.
- **Tyrimo tikslas** – Sukurtas naujas algoritminis modelis kompiuterių tinklu perduodamų duomenų efektyviam srautų panašumo vertinimui ir savybių atpažinimui.
- **Uždaviniai:**
 - Sintetiniai atvejai kompiuterių tinklo srauto transliacijai.
 - Duomenų srautų, gautų asinchroninio įrašymo metu simuliacija, modeliavimas ir jų vertinimas.
 - Tinklo srauto parametrų ir duomenų modelio sukūrimas.
 - Algoritmų kūrimas incidentų šablonų atpažinimui.
 - Sukurtų algoritmų efektyvumo vertinimas ir įtaka saugumui.
 - Gautų mokslinių rezultatų taikymas realių duomenų srautų vertinime.

Iš ankstesnės publikacijos rezultatų jau žinome, kad:

- Asinchroninis įrašų lygiavimas **yra įmanomas** nustatant tinklo anomalijas, įspėjimus ir incidentus.
- Asinchroninis įrašų lygiavimas **nesutrumpina laiko** nurodytų problemų sprendimui.
- Asinchroninio įrašų lygiavimo modelio tobulinimas **yra perspektyvus naudojant skirtingus metodus**.

- Kontroliuojama, heterogeninė tinklo aplinka.
- Asinchroninis, automatizuotas tinklo srauto įrašymas tarp grėsmės aktoiaus ir taikinio mašinos.
- Sintetiniai atakų scenarijai.

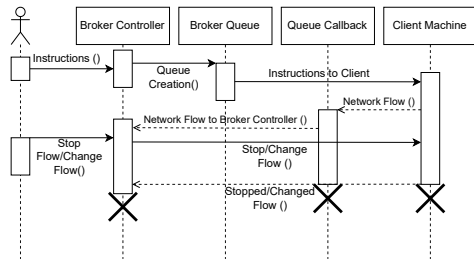


1 pav. Eksperimentinė aplinka

1 lentelė. Atributai ir vektoriai PCAP faile

Vieta paketo struktūroje	Atributas	Vektorius
TCP Payload (HTTP Request)	http.request.uri	/beacon
TCP Payload (HTTP Request)	text	//4A
TCP Payload (TLS ClientHello)	tls.handshake.ja3	78f0dc5ac5b19daf13...
TCP Header	tcp.dstport	8888
TCP Header	tcp.dstport	5000
TCP Header	tcp.flags	SYN

- Sukurta aplinka tenkina tinklo srautui keliamus reikiamus priklausomai nuo paduodamų parametrų.
- Sukurtoje aplinkoje vykdomi eksperimentai susiję su naujos algoritminės metodikos pritaikymu dideliems tingo srautams, kibernetinių atakų ir anomalijų detektavimui.



2 pav. Eksperimentinės aplinkos tyrimo objekto judėjimo kryptis

Definition (Duomenų šaltiniai)

Asinchroniniai duomenų šaltiniai iš heterogeninių taškų pateikiami kaip $\mathcal{D}_t = \{d^1, d^2, \dots, d^n\}_t$. Kur n - tai bet koks duomenų šaltinių, kurie gali būti stebimi sinchronizuotu laiku t , skaičius.

Visus tinklą stebinčius asinchroninius duomenų šaltinius apibrėžiame kaip $\mathcal{D}^*$ su bet koku laiko intervalu. Du skirtingu laiku esantys duomenų šaltiniai apibrėžiami kaip bet kurie du duomenų šaltiniai d_t^1 ir $d_{t'}^2$, kur $d_t^1, d_{t'}^2 \in \mathcal{D}^*$ tenkina sąlygą, kad $t \neq t'$. Laikas t yra laiko žyma, naudojama vietinių duomenų atsiradimui koreliuoti.

Definition (Duomenų srautas)

Apibrėžiame $\mathcal{S}_t$ kaip tinklo stebimų paketų srautą, siunčiamą iš duomenų šaltinio, kuris yra aktorius mašina. Aktorius gali būti užpuoliko arba taikinio mašinos. $\mathcal{S}$ srautas turi paketų įrašus k paketų su tinklo duomenų šaltiniu $d = d_t^i \in \mathcal{D}_t$. Tada $\mathcal{S}_t^d = \langle P_1, P_2, \dots, P_k \rangle_t^d$, kur k yra stebimo duomenų srauto paketų skaičius.

Jei duomenų srautas yra sinchronizuotas ir nenutrūkstamas laike, duomenų srautą žymėsime kaip $\mathcal{S}_{[t;\infty]}^d$ šaltinį $d = d_{[t;\infty]}^i \in \mathcal{D}^*$.

Definition (Paketų filtras)

Paketų filtras iš duomenų srauto $\mathcal{S}_t^d$ išskiria paketus, atspindinčius sudėtingus kibernetinių atakų modelius. Tegul ϕ yra loginis predikatas, apibrėžtas paketo savybėms. Apibrėžiame paketų filtrą $\mathbb{F}$ kaip funkciją $\mathbb{F} : (\mathcal{S}_t^d, \phi) \mapsto \mathcal{S}_t^{d'}$. Čia $\mathcal{S}_t^{d'}$ yra filtruotas duomenų srautas, kuriame yra tik tie paketai, kurių savybių vertės atitinka ϕ .

Definition (Paketo savybės vektorius)

Turint duomenų srautą S_t^d , siekiama apibrėžti unikalias paketų savybių vertes kiekvienam paketui P_t^d . Tegul $\mathcal{F} = \langle f_1, f_2, \dots, f_m \rangle$ yra paketo savybės, kur f yra P_t^d savybė. Kiekvienas P_t^d paketas S_t^d yra susijęs su baigtiniu savybių reikšmių rinkiniu. Šias savybių vertes žymime:

$$\mathcal{V}(P_t^d) = \langle p_{t1}^d, p_{t2}^d, \dots, p_{tj}^d \rangle$$

Kur j yra paketų savybių reikšmių skaičius sraute S_t^d .

Tegul $\mathcal{S}_t^d$ yra paketų srautas su savybėmis P_k ir atvykimo laiku $T(P_k)$, o $\mathbb{F} : (\mathcal{S}_t^d, \phi) \mapsto \mathcal{S}_t^{d'}$ sukuriantis filtruotą srautą $\mathcal{S}_t^{d'} = \{ P_k \in \mathcal{S}_t^d \mid \phi(P_k) = \text{true} \}$.

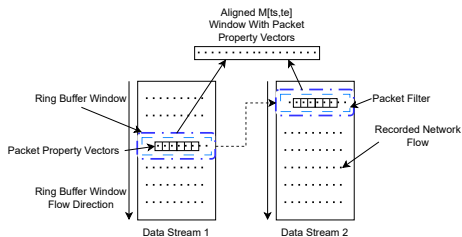
Definition (Išlygiuotas atitikimų laike langas)

Duotas lango dydis $\omega \in \mathbb{N}$ ir laikai $t_s < t_e$, šį langą apibrėžiame kaip $\mathcal{M}_{[t_s, t_e]}$, kaip filtruotų duomenų rinkinį filtruotame sraute:

$$\mathcal{M}_{[t_s, t_e]} = \langle P_k \in \mathcal{S}_t^{d'} \mid T(P_k) \in [t_s, t_e] \wedge \|\langle P_j \rangle\| \leq \omega \rangle \quad (1)$$

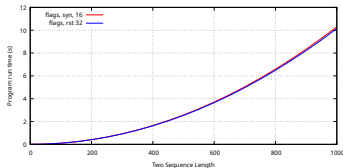
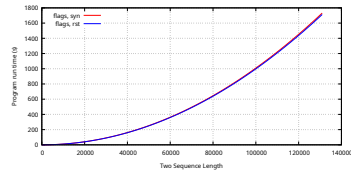
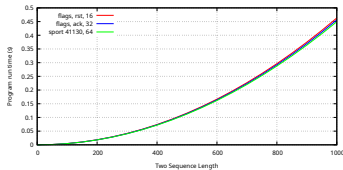
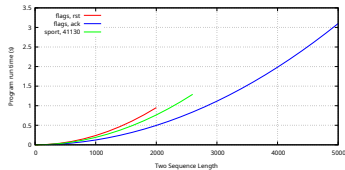
$\mathcal{M}_{[t_s, t_e]}$ yra tiksliai tie paketai, kurie atitinka ϕ kriterijų ir atkeliauja per $[t_s, t_e]$ daugiausia ω išgautų paketų.

- Pasinaudojus sukurta aplinka gaunami tinklo srautai kurie įrašomi į PCAP (angl. Packet Capture) failus. Taip pat naudojami srautai iš viešai pasiekiamų bibliotekų internete.
- Gauti objektai pritaikomi modeliui kibernetinių atakų aptikimui mažuose ir dideliuose tinklo srautuose.
- Modelyje naudojami slankūs langai su fiksuotu žiediniu buferiu kuriame vyksta paketu filtravimas pritaikant mūsų įvedamus atakų šablonų požymius.



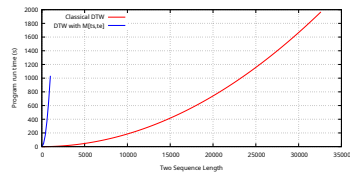
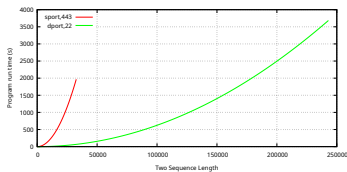
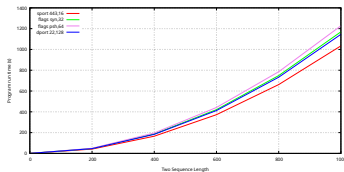
3 pav. Paralelinis PCAP failų apdorojimas su slenkančiais langais ir žiediniu buferiu

Rezultatai su mažais duomenų srautais generuotais pasinaudojus eksperimentine aplinka. Matuojamas efektyvumas tarp klasikinio DTW ir mūsų kuriamo modelio.



4 pav. PCAP faila su NMAP prievadų ir paketų vėliavų atakomis tarp atakuotojo ir taikinio virtualių mašinų

Efektyvumas tarp didelių srautų pasinaudojus klasikinio DTW ir mūsų siūlomo modelio atveju.



5 pav. PCAP objektai iš didelių srautų prieinami interneto šaltiniuose

Kokios slenkančio lango savybės padeda DTW veikti greičiau:

- Fiksuota lango trukmė. Ribojamas sekos ilgis, DTW matrica pasiekia tik nustatytas ribas.
- Fiksuotas baitų limitas vienam langui. Langai neapkraunami dideliais duomenų kiekiais.
- Žingsniavimas per langus. Gauname tankius, bet valdome kiekio langus, o ne nereguliaras milžiniškas atkarpas.
- Vektorių filtracija. Vietoje $n \times m$ „viskas su viskuo“ matricos DTW vykdoma tik ten kur reali prasmė.
- Lango atnaujinimas. Stumdant langą per srautą pasišalina pasene paketai neperkraunant viso lango iš naujo. Mažesnės sąnaudos ruošiant kandidatuojančius langus DTW.

- DTW su langais pritaikymas yra greitesnis išlygiuojant duomenis tarp skirtingų sekų į prasmingą langą.
- Greitis kyla dėl duomenų sumažinimo per langus kur DTW yra $O(n^2)$ problema mažinama iki $O(n \cdot m) \rightarrow O(n' \cdot m')$, kai $n', m' \ll n, m$
- Technologijos pasirinkimas turės įtakos efektyvumo vertinimui.

Doktorantūros studijų planas



Studijų metai	Egzaminai	
	Planas	Įvykdyta
I (2023/2024)	1	1
II (2024/2025)	2	2
III (2025/2026)	1	1
IV (2026/2027)		

2 lentelė. Egzaminai pagal studijų planą

Egzaminai 2025-2026 I pusmetis		
Planas	Įvykdyta	Būklė
Tinklų modeliai ir algoritmai (2026-01-09)	Tinklų modeliai ir algoritmai (2026-01-09)	Išlaikytas

3 lentelė. Išlaikyti egzaminai

Studijų metai	Dalyvavimas konferencijose				Publikacijos					
	Tarptautinėse		Nacionalinėse		Su citav. rodikliu			Be citavimo rodiklio		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė	Planas	Įvykdyta	Būklė
I (2023/2024)	1	1						1	1	Publikuota
II (2024/2025)					1	0	Įteikta			
III (2025/2026)					1					
IV (2026/2027)	1							1		
Iš viso:	2	1			2			2	1	

4 lentelė. Konferencijų ir publikacijų planas