Įrodymais grįsto tinklo srauto duomenų modeliavimas ir analizė kibernetinių incidentų užkardymui

Ataskaita už 2023-2024 mokslo metus. Doktorantūros pradžios ir pabaigos metai: 2023-2027

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

Vilniaus Universitetas
Matematikos ir Informatikos Fakultetas
Informatikos Institutas
Kibernetinio Saugumo Laboratorija

2025-03-28

Problema, Hipotezė

- **Problema:** Dabartiniai tinklo srauto analizės metodai yra neefektyvūs dėl didelių skaičiavimo resursų poreikio ir lėto veikimo, ypač kai reikia apdoroti didelės apimties duomenis. Realiam tinklo saugumo užtikrinimui reikalingi veiksmingesni algoritmai ir realaus laiko analizės sprendimai.
- **Hipotezė:** Tinklo srauto analizė yra neefektyvi dėl esamų algoritmų lėto veikimo, kuris kyla dėl didelių resursų reikalavimų apdorojant didelius duomenų kiekius.

Tyrimo objektas, tikslas, uždaviniai

- **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.

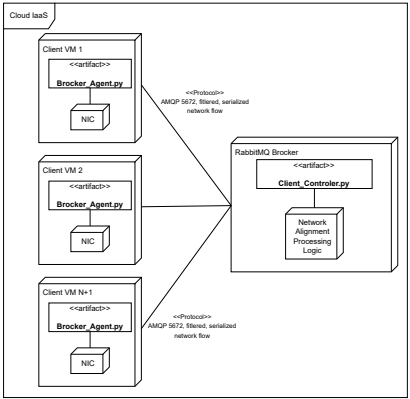
Mokslinių rezultatų pristatymas

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**.

Tyrimo metodas / aplinka

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



1 pav. Eksperimentinė aplinka

Definition (Duomenų šaltiniai)

Asinchroniniai duomenų šaltiniai iš heterogeninių taškų pateikiami kaip $\mathcal{D}_t = \{d^1, d^2, \ldots, d^n\}_t$. 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 kokiu 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.

Formalus apibrėžimas II

Definition (Duomenų srautas)

Apibrėžiame $\mathcal{S}_t$ kaip tinklo stebimų paketų srautą, siunčiamą iš duomenų šaltinio, kuris yra aktoriaus mašina. Aktorius gali būti užpuoliko arba taikinio mašinos. S srautas turi paketų įrašus k paketų su tinklo duomenų šaltiniu $d = d_t^i \in \mathcal{D}_t$. Tada $\mathcal{S}_t^d = \langle P_1, P_2, \ldots, P_k \rangle_t^d$, 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 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 $\mathbb{F} : \left(\mathcal{S}_t^d,\ \phi \right) \mapsto \mathcal{S}_t^{d\prime}$. Čia $\mathcal{S}_t^{d\prime}$ yra filtruotas duomenų srautas, kuriame yra tik tie paketai, kurių savybių vertės atitinka ϕ.

Definition (Paketo savybės vertė)

Turint duomenų srautą $\mathcal{S}_t^d$, siekiama apibrėžti unikalias paketų savybių vertes kiekvienam paketui P_t^d. Tegul $\mathcal{F} = \langle f_1, f_2, \ldots, f_m \rangle$ 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:

$$\mathcal{V}(P_t^i) = \langle p_{t\,1}^d, p_{t\,2}^d, \ldots, p_{t\,j}^d \rangle$$

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

Formalus apibrėžimas IV

Slenkamasis langas yra fiksuoto dydžio langas, kuris juda per duomenų seką, kad būtų galima apdoroti duomenų poaibius. Jis paprastai naudojamas srautinių duomenų analizėje, kad būtų galima apdoroti nepertraukiamus įvestus duomenis.

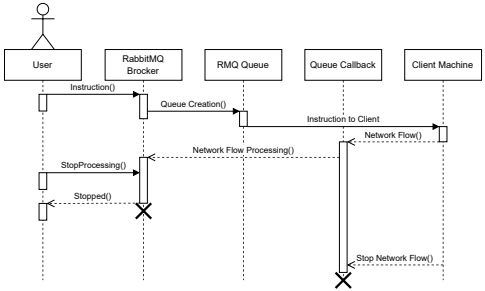
Definition (Slankantis langas)

Apibrėžiame lango dydį ω. Laiku pagrįstą slankiojantį langą virš S_t^d apibrėžia pradžios laikas t_s ir pabaigos laikas t_e, kai $t_s < t_e$. Šį langą žymime kaip

$$\mathcal{W}_{[ts,te]} = P_k \in \mathcal{S}_t^d \mid T(P_k) \in [t_s, t_e]$$

Aplinka algoritmams ir eksperimentams

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



2 pav. Eksperimentinės aplinkos tyriamojo objekto judėjimo kryptis

2024/2025 m. m. I pusmečio atlikti darbai

1. Išlaikytas privalomas dalykas „Fundamentalieji informatikos ir informatikos inžinerijos metodai“.
2. Sukurtos aplinkos tobulinimas būsimiems eksperimentams atlikti.
3. Dalinis modelio įgyvendinimas incidentų atpažinimui ir sudėtingų kibernetinių atakų aptikimui dideliuose tinklo srautuose.

2024/2025 m. m. II pusmečio suplanuoti ir vykdomi darbai

1. Išlaikyti pasirenkamą dalyką „Šiuolaikinės duomenų bazių sistemos“.
2. Mokslinė publikacija „Computes and Security“ žurnale.
3. Dalyvavimas „Cybersecurity and Privacy (CySeP)“ vasaros mokykloje.

Doktorantūros studijų planas

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

1 lentelė. Egzaminai pagal studijų planą

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	Priimta
II (2024/2025)					1					
III (2025/2026)					1					
IV (2026/2027)	1							1		
Iš viso:	2	1			2			2	1	

2 lentelė. Konferencijų ir publikacijų planas

Ataskaitinis studijų metai (2024/2025)

Dalyvavimas tarptautinėje konferencijoje 2023/2024 (II Pusmetis)		
Planas	**Įvykdyta**	**Konferencijos tipas**
23rd European Conference on Cyber Warfare and Security - ECCWS 2024	Virgilijus Krinickij and Linas Bukauskas, Asynchronous Record Alignment of Network Flows for Incident Detection and Reconstruction, 23rd European Conference on Cyber Warfare and Security - ECCWS 2024, Agora Building at University of Jyväskylä, Finland 26 - 28 June 2024	Tarptautinė

3 lentelė. Dalyvavimas tarptautinėje konferencijoje

Egzaminai 2024-2025 I pusmetis		
Planas	**Įvykdyta**	**Būklė**
Fundamentalieji informatikos ir informatikos inžinerijos metodai (2025-01-28)	Fundamentalieji informatikos ir informatikos inžinerijos metodai (2025-01-28)	Išlaikytas

4 lentelė. Išlaikyti egzaminai

Ataskaitinis studijų metai (2024/2025) (2)

Publikacijos 2023/2024 (II pusmetis)			
Planas	**Įvykdyta**	**Būklė**	**Publikacijos Tipas**
23rd European Conference on Cyber Warfare and Security - ECCWS 2024, Agora Building at University of Jyväskylä, Finland 26 - 28 June 2024	Krinickij, Virgilijus; Bukauskas, Linas. Asynchronous record alignment of network flows for incident detection and reconstruction // European Conference on Cyber Warfare and Security: Proceedings of the 23rd European Conference on Cyber Warfare and Security. Reading: Academic Conferences International Limited. ISSN 2048-8602. eISSN 2048-8610. 2024, vol. 23, no. 1, p. 249-256. DOI: 10.34190/eccws.23.1.2254.	Publikuota	Be cituojamumo rodiklio

5 lentelė. Publikacijos

Publikacijos (tik su citavimo rodikliu)		
	Bibliografinis aprašas	**Būklė**
1.		Ruošiama

6 lentelė. Publikacijos su cit. rodikliu