LIETUVOS GELEŽINKELIAI

Kaip generatyvinis dirbtinis intelektas (GenDI) pakeitė kibernetinių incidentų tyrimus?

Dr. Šarūnas Grigaliūnas - Lietuvos Geležinkelių skaitmeninės saugos vadovas

Dr. Šarūnas Grigaliūnas

Šarūnas yra ISECOM sertifikuotas įsiskverbimo testuotojas, sertifikuotas ISO 27001 ISMS auditorius.

Kauno technologijos universitete, įgijo daktaro laipsnį.

Daugiau nei 16 metų patirtis IT audito ir saugumo srityje, domina kibernetinis saugumas. 5 metus dirbo vyresniuoju kibernetinio saugumo konsultantu įmonėje.

PATIRTIS

Lietuvos geležinkeliai

Šarūnas buvo atsakingas už sistemų stebėseną ir atitiktį teisės aktų reikalavimams, galutinių taškų sistemų saugumo priemonių ir taikomųjų programų valdymą ir diegimą, tinklo ir ugniasienės saugumo valdymą, informacinių sistemų saugumo pažeidžiamumų valdymą, informacijos saugumo politikos kūrimą ir geležinkelių SCADA saugumo pažeidžiamumų valdymą.

ISACA LT

Užtikrinti lyderystę svarbiausiose srityse, tokiose kaip saugos vertinimas, testavimas, specifikacijos, informacijos saugumo valdymo sistemos, tapatybės valdymo ir privatumo technologijos.

Transcendent Group

"Transcendent Group Lietuva" Šarūnas yra informacinių ir ryšių technologijų praktikos vadovas, daugiausia dėmesio skiriantis atitikties kibernetinio saugumo paslaugų plėtrai.

LITNET CERT

Šarūnas sukūrė integruotą strateginės, operatyvinės ir taktinės kibernetinio saugumo grėsmių žvalgybos modelį. Slaptų kibernetinių atakų ir naujų kenkėjiškų programų aptikimas, ankstyvas įspėjimas ir selektyvus grėsmių duomenų platinimas - tai tik keletas techninės grėsmių žvalgybos privalumų. Sprendimų priėmimo procesas (apsauga ir tobulinimas).

Kauno Technologijos Universitetas

Šarūnas yra socialinių tinklų saugumo, kibernetinių atakų, skaitmeninės kriminalistikos įrodymų psichologijos ir praktinio pritaikymo, naujų žaidybinimo metodų, skirtų saugumo sąmoningumui ugdyti, saugumo kultūrai kurti ar vidinėms grėsmėms valdyti, analitikas.

FOKUSAS

CERT / GEANT / LitNET

- Kibernetinių grėsmių žvalgyba ir analizė, saugumo incidentų nagrinėjimas, skaitmeninė ir kibernetinių nusikaltimų ekspertizė
- MTEP, vidaus sistemų ir (arba) paslaugų pažeidžiamumo vertinimas, vadovavimas serverių apsaugai
- WP8 Task1 Saugumo valdymas

Saugumo valdymas 2017- iki dabar

- Operacinių technologijų ir telekomunikacijų saugumo auditas, įsiskverbimo testavimas
- Pramonės valdymo sistemų/ SCADA saugumo valdymas
- Nacionalinės ypatingos svarbos infrastruktūros dalies saugumo užtikrinimas
- Teisinė ir reguliavimo atitiktis

Pratybos 2016 - 2023

- Kibernetinės gynybos pratybos, struktūrizuoti žvalgybos analizės metodai ir grėsmių modeliavimas.
- kritinės infrastruktūros ir prietaisų testavimas.
- Mokymai (bendrieji, saugaus kūrimo, pentest, socialinės inžinerijos, skaitmeninės kriminalistikos ir kt.).

SERTIFIKATAI

ISO 27001- Information Security Management System Auditor

OPST - Certified OSSTMM 3.0 professional security tester

OPSA - Certified OSSTMM 3.0 professional security analyst

CTA - Certified Trust Analyst from ISECOM (Institute for Security and Open Methodologies

ICS – ICS-CERT U.S. Department of Homeland Security, Control Systems Cybersecurity

ACE - AccessData Certified Examiner

01

Kas yra kibernetinis incidentas?

Kibernetinis incidentas apibrėžiamas kaip **įvykis arba veika kibernetinėje erdvėje**, kurie gali sukelti arba **sukelia grėsmę** arba neigiamą poveikį ryšių ir informacinėmis sistemomis perduodamos ar jose tvarkomos **elektroninės informacijos prieinamumui, autentiškumui, vientisumui ir konfidencialumui**, taip pat gali trikdyti **arba trikdo ryšių ir informacinių sistemų veikimą, valdymą** ir paslaugų jomis **teikimą**.

02

Generatyvinis dirbtinis intelektas (GenDI)

GenDI

Generatyvinis dirbtinis intelektas - sistemos ir algoritmai, gebantys savarankiškai generuoti turinį – tekstą, vaizdus ar kitų formų duomenis

Egzistuojančios grėsmės

- 2024 m. "OpenAI" teigia, kad paslaptingos pokalbių istorijos atsirado dėl paskyros perėmimo.
- 2023 m. rugpjūčio mėn. įsilaužėlių grupė pareiškė, kad pažeidė ChatGPT duomenų bazę ir nutekino daugiau nei 10 milijonų naudotojų, kurie bendravo su pokalbių robotu, pokalbių žurnalų. Pokalbių žurnaluose buvo slaptos ir privačios informacijos, tokios kaip vardai, adresai, telefonų numeriai, seksualinės nuostatos, sveikatos būklė ir kt.
- 2023 m. birželio mėn. buvo pranešta, kad nuo 2022 m. birželio mėn. iki 2023 m. gegužės mėn. buvo pažeista daugiau nei 100 000 „ChatGPT“ paskyros profilių ir jie buvo parduoti tamsiosiose žiniatinklio prekyvietėse (DarkNet).

O kaip LTG?

- Dokumentas, įtvirtinantis principus, kuriais remiantis būtų užtikrintas teisėtas, vienodas, nuoseklus, saugus ir atitinkantis reglamentavimą GenDI sistemų naudojimas LTG grupėje.
- Principai detalizuoja GenDI sistemų saugaus naudojimo LTG grupėje veiklą.
- Nustato ir reglamentuoja bendruosius principus.
- Užtikrina vienodą ir nuoseklų GenDI sistemų valdymą, atsakomybę ir kontrolės metodus.

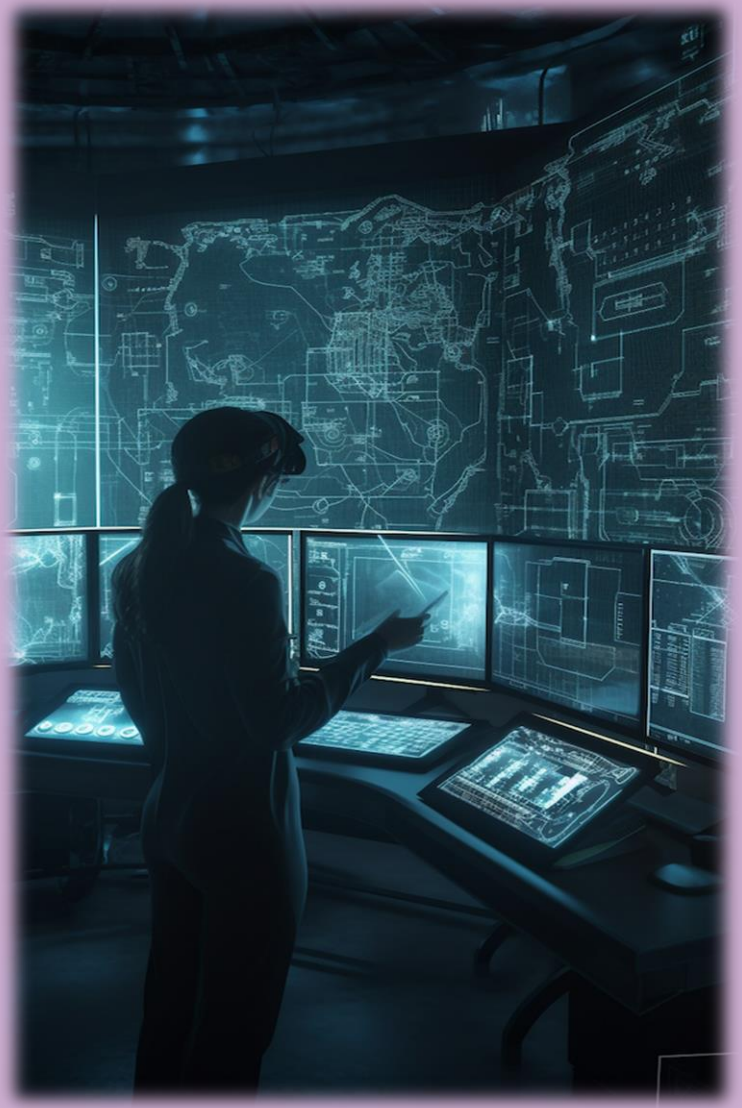
NKSC “Parengtos Generatyvinio dirbtinio intelekto (GenDI) saugaus naudojimo organizacijoje gairės” https://www.nksc.lt/naujienos/parengtos_generatyvinio_dirbtinio_intelekto_gendi_.html

GenDI sistemų saugaus naudojimo principai

- LTG grupė savo veikloje naudoja tik saugias GenDI sistemas, kurios įtrauktos į leistinų naudoti LTG grupės veikloje GenDI sistemų sąrašą – „White List“
- Draudžiama naudoti GenDI sistemas ypatingos svarbos informacinėje infrastruktūroje (YSII).
- LTG grupė nustato pareigybių, kurioms leistina naudoti GenDI sistemas (kurioms suteikiami prisijungimo prie GenDI sistemų duomenys), sąrašą.
- Panaudojus GenDI sistemas gautas rezultatas žymimas kaip GenDI sugeneruotas atsakymas.

03

Problema



Žmonės

- Brangu samdyti aukštos klasės specialistą.
- Kibernetinės saugos kompromisai.
- Atakuotojai turi didesnius biudžetus, naudoja naujausias technologijas

Segmentavimas

- Tinklo perimetras.
- Informacinės sistemos (paslaugų)
- Galiniai naudotojai

Incidentai

- Reakcijos laikas
- Tyrimo trukmė
- Šaltiniai (žurnalai, tinklo srautai, svetainių turinys)

GenDI privalumai ir ribojimai

Privalumai

- Greičiau identifikuoja ir analizuoja incidentus.
- Automatizuoja ataskaitų kūrimą ir rekomendacijų pateikimą.
- Gali padėti SOC (IT) komandai tobulėti, mokydamasis iš ankstesnių incidentų.

Ribojimai

- Kartais gali pateikti neteisingą ar nepilną informaciją.
- Yra priklausoma nuo mokymo duomenų kokybės ir gali reikėti papildomos žmogiškosios priežiūros.

Panaudojimo atvejai

Informacinių pranėšimų (alerts) tyrimas

Duomenų srauto analizė

Ataskaitos

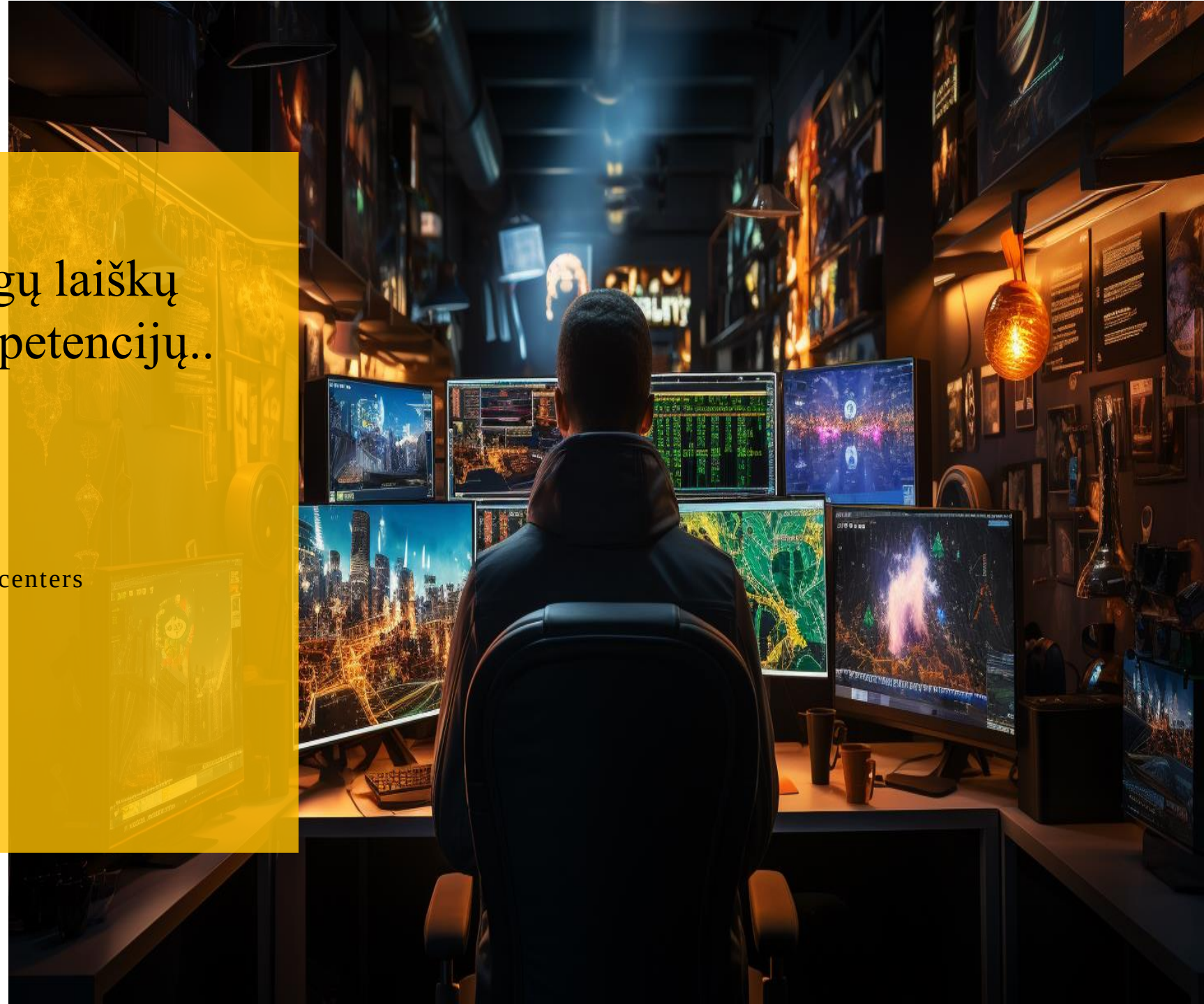
04

Informacinių pranėšimų tyrimas

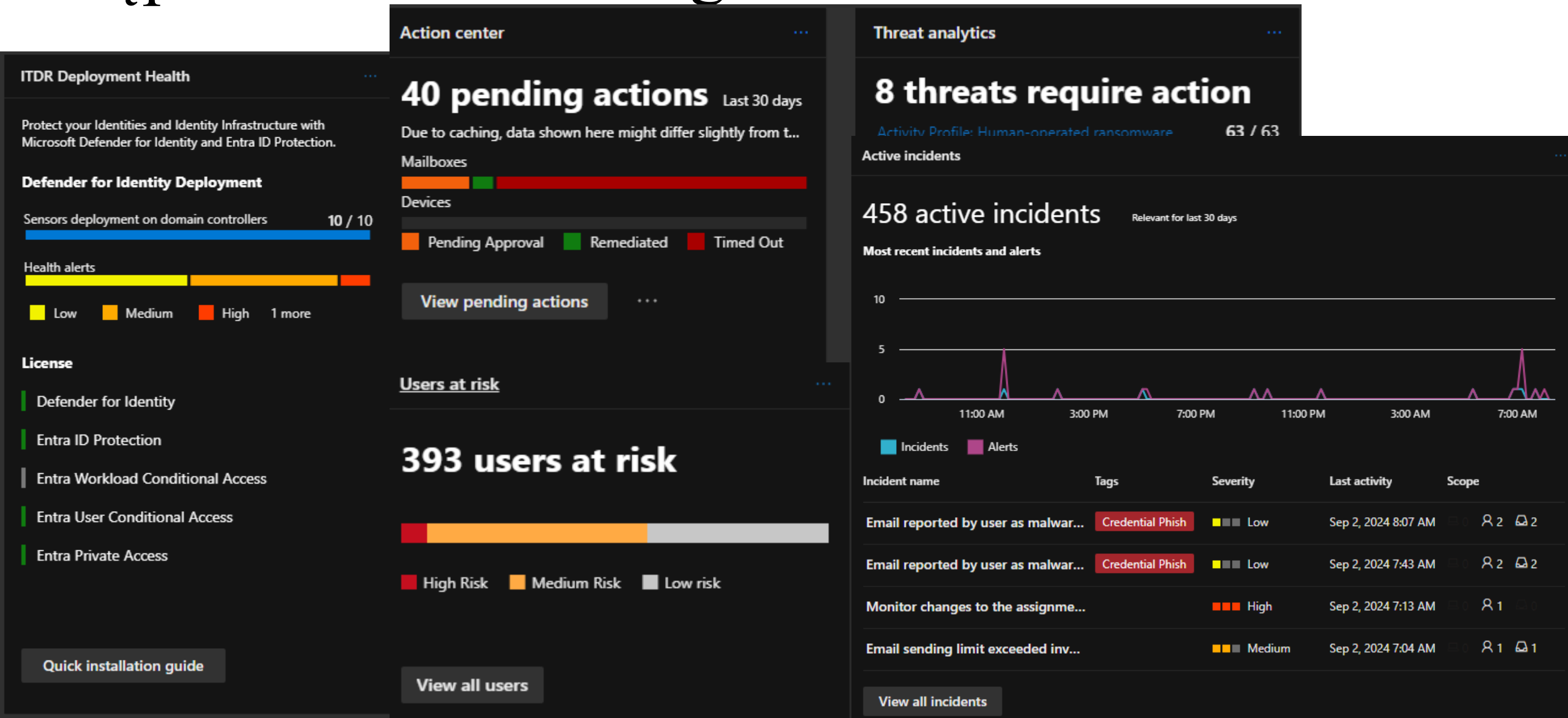
Reikia ištirti apgaulingų laiškų ataką, bet trūkstą kompetencijų..

ĮRANKIAI (AI)

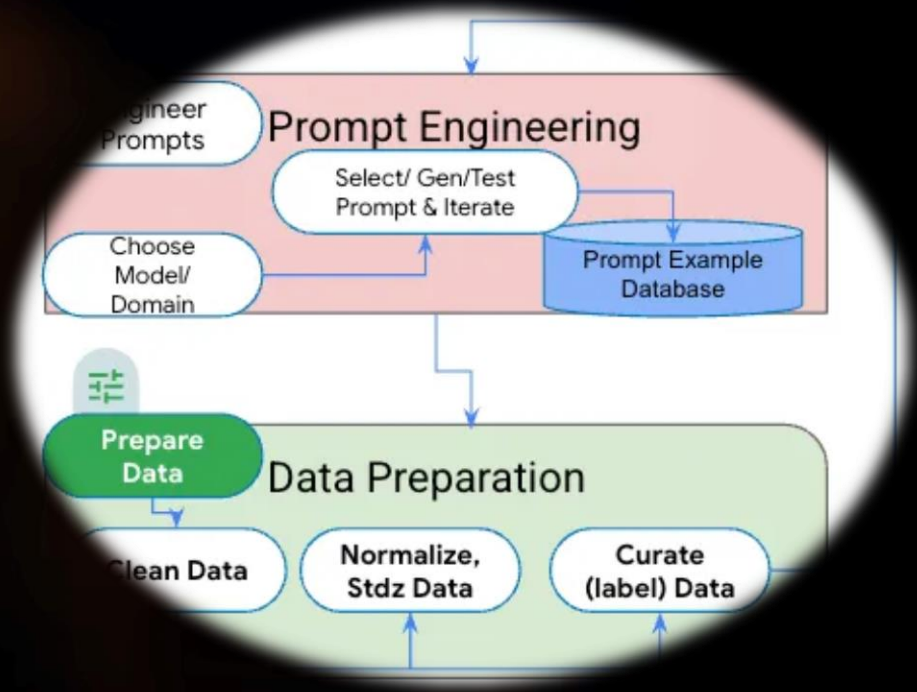
Microsoft security portals and admin centers
chatgpt-retrieval
PentestGPT
Cloude AI



Įprastas vaizdas saugos komandai

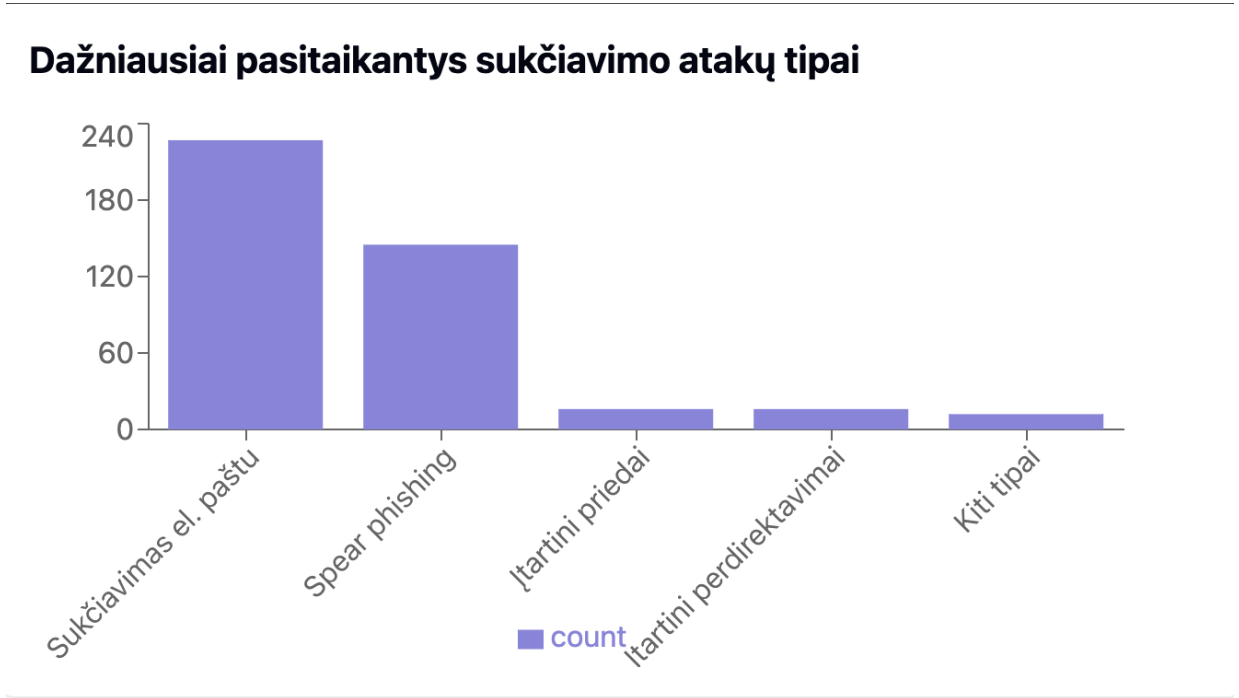


O jei GenDI kvieti į pagalbą?



Šaltinis: SUBSERVIENCE Official Trailer (2024)

Reikalinga atlikti pirminę analizę, nustatyti dėsningumus, tendencijas ir pateikti rekomendacijas

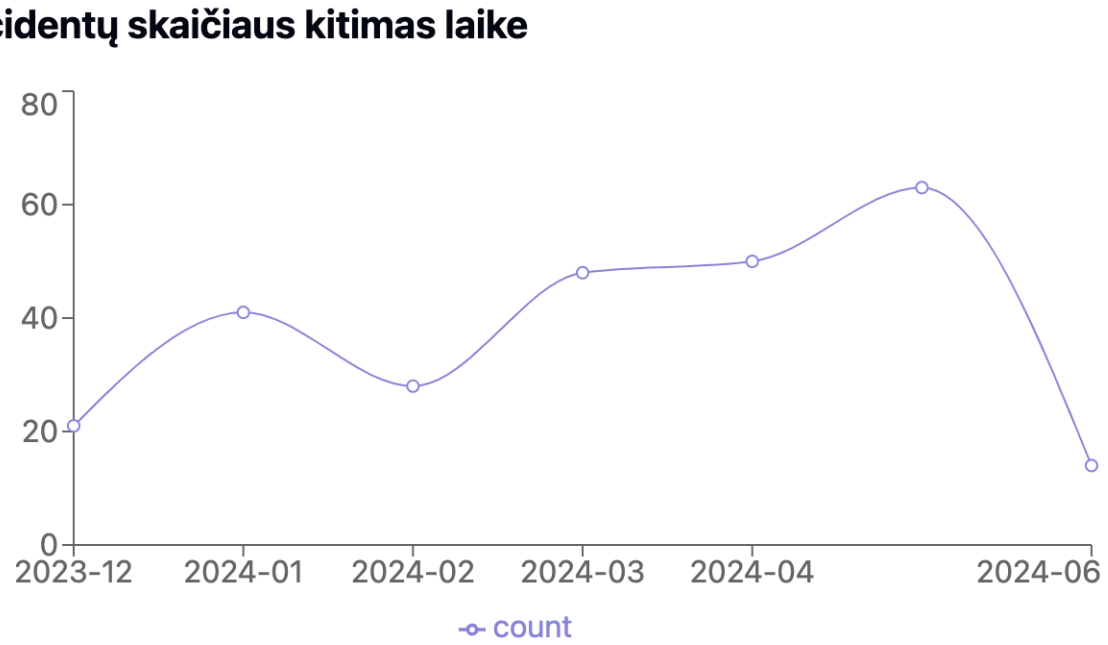


I dalis Dažniausiai pasitaikantys sukčiavimo atakų tipai

Analizė:

- Sukčiavimas el. paštu yra dažniausiai pasitaikantis atakų tipas su 227 užfiksuotais atvejais.
- Antroje vietoje yra spear phishing atakos su 145 atvejais.
- Įtartini priedai ir įtartini perdirektavimai yra rečiau pasitaikantys, bet vis dar reikšmingi su 16 atvejų kiekvienas.
- Kiti tipai sudaro mažesnę dalį su 12 atvejų.

Reikalinga atlikti pirminę analizę, nustatyti dėsningumus, tendencijas ir pateikti rekomendacijas

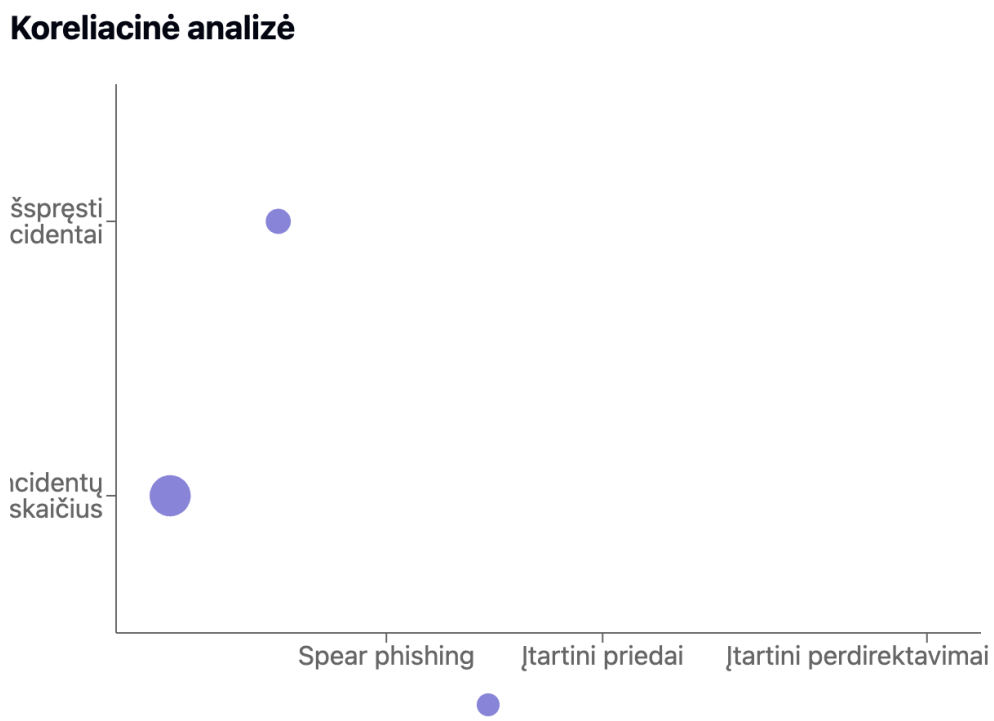


II dalis Laikotarpiai, kai incidentų pasitaiko dažniausiai

Analizė:

- Incidentų skaičius rodo augimo tendenciją nuo 2023 m. gruodžio iki 2024 m. gegužės.
- Didžiausias incidentų skaičius užfiksuotas 2024 m. gegužės mėnesį (63 atvejai).
- Mažiausias incidentų skaičius buvo 2023 m. gruodžio mėnesį (21 atvejis).
- 2024 m. birželio mėnesio duomenys gali būti neišsamūs, nes apima tik dalį mėnesio.

Reikalinga atlikti pirminę analizę, nustatyti dėsningumus, tendencijas ir pateikti rekomendacijas



III dalis Koreliacinė analizė

Analizė:

- Stipriausia teigiama koreliacija (0.8) yra tarp sukčiavimo el. paštu ir bendro incidentų skaičiaus.
- Spear phishing taip pat rodo stiprią koreliaciją (0.7) su bendru incidentų skaičiumi.
- Įtartini priedai ir perdirektavimai rodo vidutinę koreliaciją (0.4) su incidentų skaičiumi.
- Išspręstų incidentų koreliacija su atakų tipais yra silpnesnė, o tai gali rodyti, kad sprendimo strategijos nėra pakankamai efektyvios arba pritaikytos konkretiems atakų tipams.

Reikalinga atlikti pirminę analizę, nustatyti dėsningumus, tendencijas ir pateikti rekomendacijas

IV dalis Remiantis duomenų analize, štai 3 dažniausiai pasitaikantys ir pasikartojantys sukčiavimo incidentų tipai:

1. Įtartina nuoroda vidiniame tikslinės atakos el. laiške (82 incidentai) Šio tipo sukčiavimas apima el. laiškus, siunčiamus iš organizacijos vidaus (arba atrodančius, kad yra iš vidinių šaltinių), kuriuose yra įtartinos nuorodos. Tai dažniausias atakos tipas, rodantis reikšmingą grėsmę iš potencialiai pažeistų vidinių paskyrų arba sudėtingų apsimetimo metodų.
2. Galima tikslinė ataka per nuorodą (77 incidentai) Tai tiksliniai sukčiavimo bandymai naudojant kenkėjiškas nuorodas, greičiausiai pritaikyti konkretiems asmenims ar grupėms organizacijoje. Didelis šio tipo dažnumas rodo, kad atakuotojai deda nemažai pastangų kurdami personalizuotus, įtikinamus sukčiavimo bandymus.
3. Tikslinės atakos el. laiškas su žinomais sukčiavimo požymiais (68 incidentai) Šie incidentai apima el. laiškus, kurie turi charakteristikas, dažnai siejamas su sukčiavimo bandymais. Faktas, kad šie yra "žinomi" požymiai, rodo, kad nors saugumo sistemos gali juos aptikti, jie vis dar išlieka paplitusia grėsme.

05

Duomenų srauto analizė

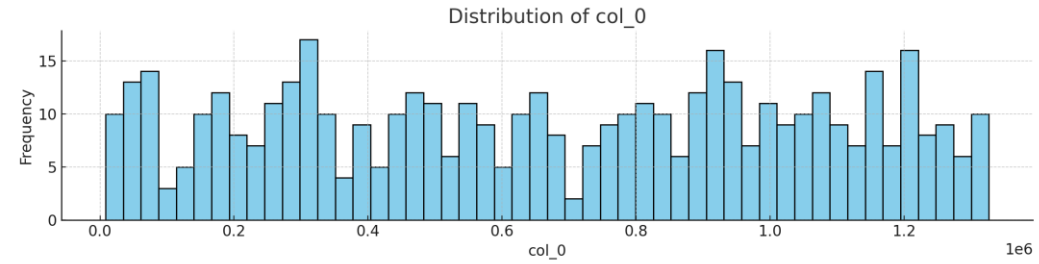
Neturite analitiko?

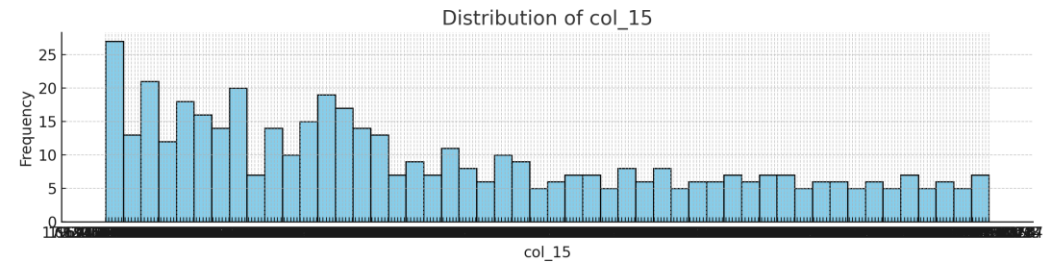
ĮRANKIAI (AI)

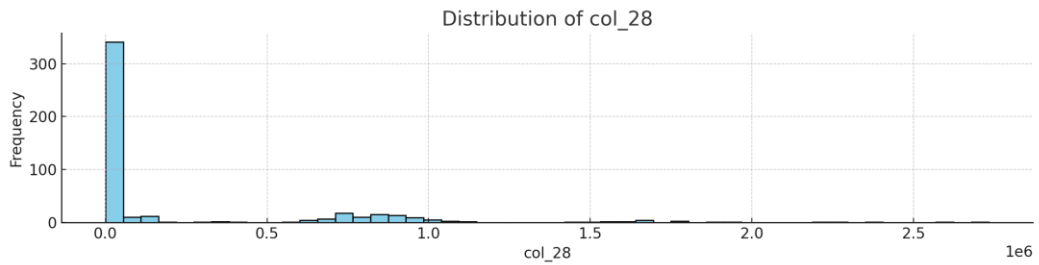
ChatGPT
ElevenLabs
DeepL

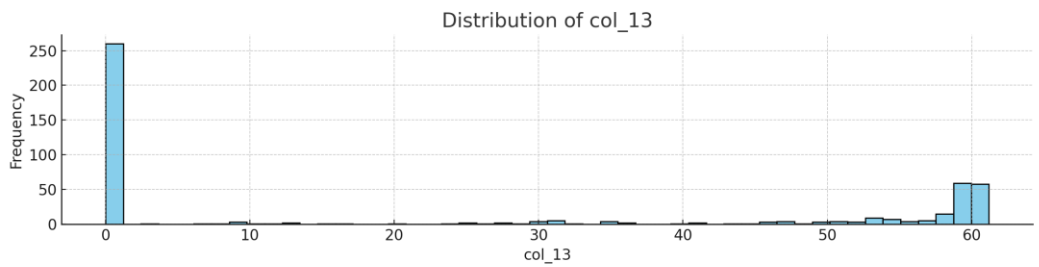
TYRIMO ĮRANKIAI

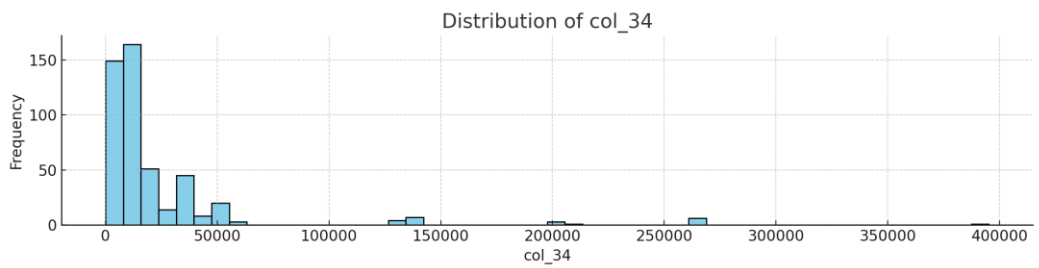
Neteble
Gemini
OpenAI Playground











Pirmas žingsnis

Norint nustatyti kibernetinės atakos tipą, analizuojant tinklo srauto duomenis reikia įvertinti įvairius duomenų modelius, elgseną ir charakteristikas. Kai kurie bendri įvairių tipų kibernetinių atakų rodikliai ar modeliai yra šie:

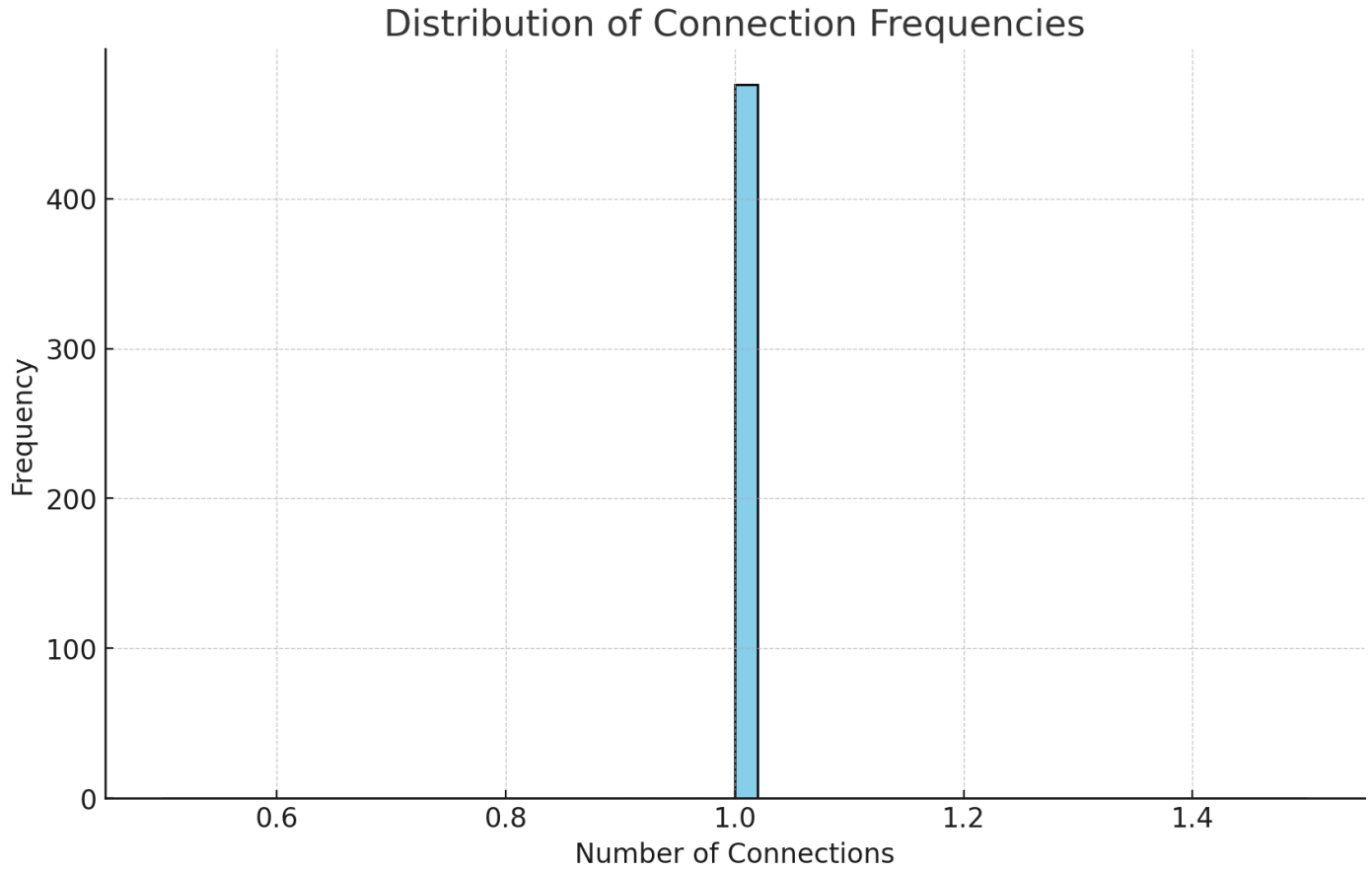
1.Prievadų skenavimas: Dažni prisijungimai prie kelių prievadų, kai duomenų perduodama nedaug arba visai neperduodama.

2.DDoS (angl. Distributed Denial of Service): Didelis užklausų ar prisijungimų skaičius per trumpą laiką, dažnai iš kelių šaltinių.

3.Duomenų eksfiltracija (angl. Data Exfiltration): Nuolatinis ir didelis duomenų perdavimas, ypač į išorinius arba neįprastus IP adresus.

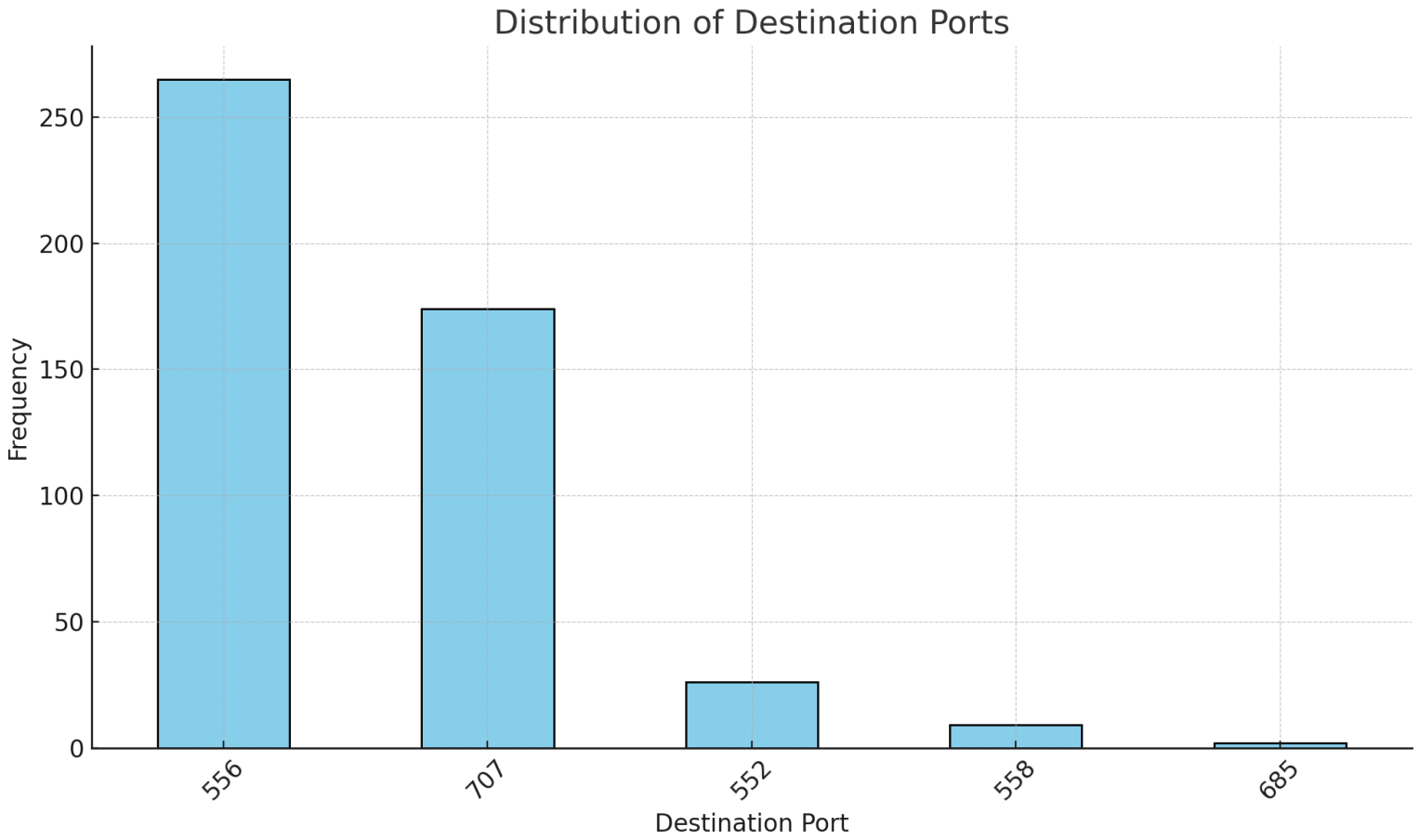
4.Brute force ataka: Atvirkštinis įsilaužimas: pasikartojantys bandymai prisijungti arba prisijungimai iš vieno IP arba kelių IP adresų.

Antras žingsnis



Iš pasiskirstymo matyti, kad dauguma jungčių atsiranda tik vieną kartą. Šis modelis nėra būdingas DDoS atakai, nes tokio tipo atakos atveju galėtume tikėtis dažnesnių pasikartojančių prisijungimų.

Trečias žingsnis



Atsižvelgiant į pastebėtus dėsningumus:

- Duomenys nerodo nuoseklių didelių duomenų perdavimų, kurie leistų daryti prielaidą apie duomenų nutekėjimą.
- Nėra aiškių skenavimo ar DDoS požymių.

Ketvirtas žingsnis

Atsiųstame paveikslėlyje parodytas paskirties prievadų pasiskirstymas "NetFlow" duomenų rinkinyje. Du dažniausiai pasitaikantys paskirties prievadai yra 955 ir 707. Šie prievadai dažniausiai naudojami atitinkamai failų dalijimosi ir nuotolinio darbalaukio protokolo (RDP) srautams. Tai rodo, kad užpuolikas bando pavogti failus arba gauti prieigą prie nutolusios sistemos.
Štai keletas galimų kibernetinių atakų, kurios gali būti vykdomos remiantis paveikslėliu:

- Failų dalijimosi ataka: Užpuolikas gali naudoti failų dalijimosi paslaugą, pavyzdžiui, "BitTorrent", kad į aukos sistemą atsisiųstų kenkėjiškus failus. Šie failai gali būti naudojami duomenims vogti, kenkėjiškai programinei įrangai įdiegti arba sistemai perimti kontrolę.
- RDP ataka: Užpuolikas gali naudoti RDP, kad nuotoliniu būdu prisijungtų prie aukos sistemos. Gavęs prieigą, jis gali pavogti duomenis, įdiegti kenkėjišką programinę įrangą arba perimti sistemos kontrolę.
- Botneto ataka: Užpuolikas gali naudoti aukos sistemą kaip botneto dalį. Botnetas - tai užkrėstų kompiuterių tinklas, kurį gali valdyti užpuolikas. Botnetai gali būti naudojami nepageidaujamoms žinutėms siųsti, DDoS atakoms vykdyti arba duomenims vogti.

Analižė atlikta pagal pateiktą paskutinį (3 žingsnis) paveikslą- Google Gemini

06

Ataskaitų rengimas

Kas mėgstate rašyti ataskaitas?

ĮRANKIAI
OpenAI
ChatGPT
DeepL
Elevenlabs

Analizuojant perspėjimus (Google Gemini)

„Microsoft Defender“ sistema užfiksavo didelį kiekį įvairaus sunkumo lygio perspėjimų. Daugiausiai buvo nustatyta „Žemo“ ir „Aukšto“ sunkumo lygio incidentų. Tai rodo, kad sistema aktyviai stebi ir identifikuoja potencialius grėsmes.

Sunkumo lygis

- **Žemas (1,565)** Didžiausia dalis perspėjimų. Tai gali reikšti dažnus, bet mažiau kritinius incidentus, tokius kaip nedideli konfigūracijos klaidimai ar nesėkmingi prisijungimo bandymai iš žinomų šaltinių.
- **Aukštas (1,421)** Antra pagal dydį grupė. Šie perspėjimai signalizuoja apie potencialiai rimtesnes grėsmes, tokias kaip kenkėjiškos programos bandymai įsiskverbti į sistemą arba neįprasti tinklo aktyvumai.
- **Vidutinis (739)** Šie perspėjimai vertinami kaip vidutinio lygio rizika ir reikalauja atidžiau ištirti.
- **Informacinis (722)** Tai dažniausiai yra pranešimai apie įvykius, kurie nėra laikomi grėsme, bet gali būti naudingi stebint sistemos veikimą.

Būsena

- **Nauji (3,969)** Didžioji dalis perspėjimų dar nėra ištirta ir reikalauja dėmesio.
- **Vykdoma (478)** Šie perspėjimai yra tiriami arba jau buvo imtasi veiksmų jiems pašalinti.

GenDI informacijos saugumo asistentas

WWW.CISO.LT

Esu informacijos saugumo pareigūno asistentas Lietuvos įmonėms ir organizacijoms.

LIETUVOS GELEŽINKELIAI

Kaip generatyvinis dirbtinis intelektas (GenDI) pakeitė kibernetinių incidentų tyrimus?

Dr. Šarūnas Grigaliūnas - Lietuvos Geležinkelių skaitmeninės saugos vadovas