
JYVSECTEC
Jyväskylä Security Technology
DEFEND YOUR NETWORKS.
www.jyvsectec.fi/en

MATINE tutkimusseminaari 22.11.2018

FT Tero Kokkonen

+35840 438 5317

tero.kokkonen@jamk.fi

www.jyvsectec.fi/en

JYVSECTEC® Jyväskylä Security Technology

jamk.fi Institute of Information Technology

Hankkeen perustiedot

- Hanke: Tekoälyn käyttö poikkeamapohjaiseen tunkeutumisten havainnointiin verkkoliikenteestä
- Toteuttaja: Jyväskylän ammattikorkeakoulu, IT-instituutti
- Hankkeelle myönnetty MATINE rahoitus: 90 882€

Taustaa

- Verkotettujen tietojärjestelmien määrän kasvun, sekä kiihtyvän digitalisaation myötä yhteiskuntamme on täysin riippuvainen tietoverkoista ja niiden turvallisuudesta
- Tämä koskettaa myös Puolustusvoimia ja kaikkia yhteiskunnan turvallisuusviranomaisia, joiden (johtamis)järjestelmät ovat riippuvaisia toimivista tietoverkoista ja tietojärjestelmistä
- Erilaisten verkko- ja kyberhyökkäysten tai yleisemmin tunkeutumisten määrä kriittisiä järjestelmiä kohtaan kasvaa jatkuvasti, mikä kasvattaa tarvetta tunkeutumisten havainnointijärjestelmän kehitykselle

Taustaa

- Yleisesti termillä ”tunkeutuminen” (intrusion) tarkoitetaan tekoa tai menetelmää, jolla vaarannetaan tietojärjestelmä tai –verkko
- Verkkopohjaiset IDS –järjestelmät jaetaan kahteen eri kategoriaan niiden toiminnan mukaan;
  - poikkeamien havainnointiin perustuviin havainnointijärjestelmiin (Anomaly Based IDS) ja
  - väärinkäytösten havainnointiin perustuviin havainnointijärjestelmiin (Misuse Based IDS)

Tehtävänasettelu ja tavoitteet

- Tutkimuksessa tavoitteena oli vertailla, tutkia ja soveltaa uusimpia avoimen lähdekoodin tekoäly- / syväoppimisohjelmistokirjastoja poikkeamapohjaiseen tunkeutumisten havainnointiin verkkoliikenteestä
  - Poikkeamien havainnoinnin mukaisesti sovellukselle opetetaan normaalin/laillisen verkkoliikenteen malli ja tunnistamisvaiheessa tunnistetaan poikkeamat opitusta normaalimallista
- Tutkimushankkeen lopputuloksena rakennettiin pilottisovellus, jonka kyvykkyyttä on testattu monipuolisella hyökkäysdatalla

Tieteellinen merkittävyys

- Tekoälykehitykseen ja tutkimukseen panostetaan maailmanlaajuisesti suuria määriä rahaa ja resursseja
  - Kyberturvallisuus on nopeasti kehittyvä tekoälytutkimuksen sovellusalue

Tulosten hyödyntämismahdollisuudet

- Tutkimusten tulokset ovat sotilaallisen maanpuolustuksen ja yhteiskunnan kokonaisturvallisuuden kehittämisen kannalta erittäin tärkeitä ja sovellettavissa auttamaan kaikkien tietoverkoista riippuvien järjestelmien turvaamisessa
- Puolustushallinnon strategiset valinnat kuten verkottunut puolustusratkaisu ja toisaalta koko sodan kuvan laajeneminen kasvattavat verkon kautta tapahtuvia vaikutusmahdollisuuksia Puolustusvoimia ja koko modernia yhteiskuntaa vastaan
- Tämän vuoksi tietoverkossa tapahtuvien ilmiöiden ymmärtäminen ja johtamisen perustana tarvittava tilannekuva, sekä tilannekuvan pohjana oleva havainnointi- ja sensorikyky ovat tärkeässä asemassa
- Tutkimustuloksia voivat hyödyntää myös muut turvallisuusviranomaiset, sekä yleisemmin kaikki viranomais- ja yksityistoimijat, joiden toiminta on riippuvaista verkotetuista järjestelmistä

Tutkimusorganisaatio

- FT Tero Kokkonen, Hankkeen vastuullinen johtaja
- TkL Mika Rantonen
- Asiantuntija, FK Samir Puuska
- Asiantuntija, Ins Janne Alatalo
- Projektityöntekijä Eppu Heilimo

Tutkimuksen suoritus

- Tutkimusmetodologiana käytettiin konstruktiivista tutkimusta
  - lähtökohtana on rakentaa konstruktio, jolla pyritään vastaamaan tosielämän ongelmaan.
- Mallia testattiin RGCE Cyber Range –ympäristössä (www.jyvsectec.fi/rgce)
  - moderni ja realistinen uhkatoimijaprofiili/hyökkäysliikenne, sekä monipuoliset verkkotopologiat
  - KYHA18 –harjoituksen verkkoliikennetaltioita tutkimuksen käytössä (https://www.defmin.fi/ajankohtaista/tiedotteet?9_m=9314)

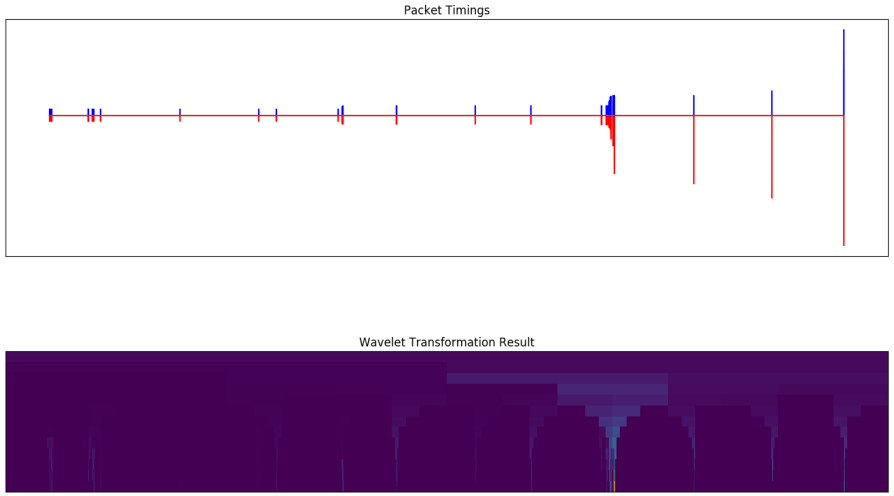
Kehitetty sovellus

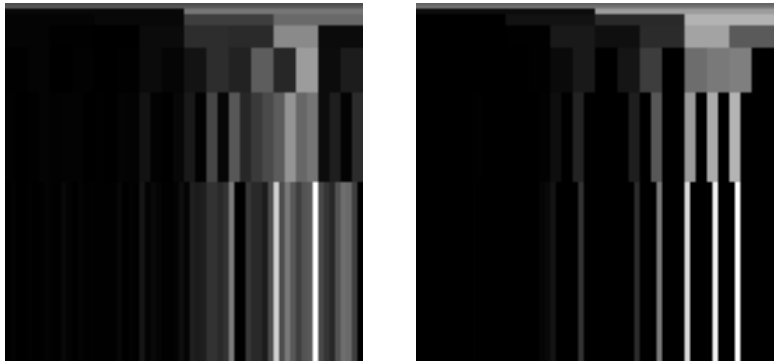
Malli

- Tavoitteena oli kyetä tunnistamaan poikkeamia TLS-salatusta Internet liikenteestä
  - Salaamaton Internet käytännössä katoamassa
  - Haittaohjelmilla salattu ja yleisesti käytetty verkkoliikenne on luonnollinen valinta komentokanavaksi
  - Salaus rajoittaa näkyvyyttä eri analyysimenetelmissä
- Tekoälytoteutuksessa käytetään Haar Wavelet muunnosta ja Adversarial Autoencoder –neuroverkkoa

Wavelet-muunnos

- Verkkoliikenne muodostaa aikasarjan salauksesta huolimatta
  - Aikasarjat ovat usein haasteellisia koneoppimiselle. Wavelet-analyysissä signaali tai kuva jaetaan approksimaatioihin sekä aika- että taajuustasossa
- Haar Wavelet skaalaa/pakkaa signaalin neliön muotoisiin sekvensseihin





(a) Legitimate web browser traffic to an authentication portal.

(b) Malicious traffic caused by a CobaltStrike beacon.

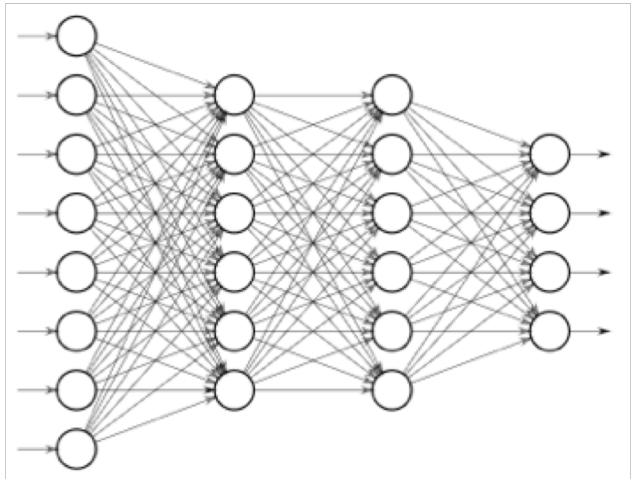
Fig. 1: Wavelet decomposition, illustrated here with scalograms, for both legitimate and malicious TLS traffic samples.

JYVSECTEC® Jyväskylä Security Technology

jamk.fi Institute of Information Technology

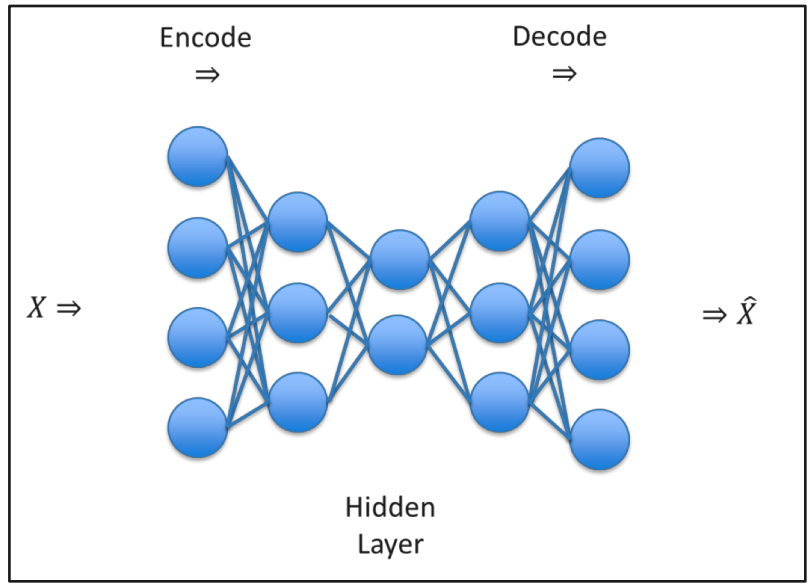
Neuroverkot

- Neuroverkot ovat verkostoja, jotka koostuvat keinotekoisista neuroneista
- Verkko laskee neuroneille syötetyn numeerisen datan annetulle pisteelle
- Opetusdatalla saadaan aikaiseksi pareja (sisäänmeno – ulostulo), joiden perusteella tunnistetaan samankaltaisesta datasta opetetut ominaisuudet
  - Pyritään oppimaan muuttujien epälineaariset riippuvuussuhteet
- Neuroverkolla on sisään- ja ulostulokerrokset sekä väliin jäävät piilotetut kerrokset. Piilotetuissa kerroksissa tapahtuu ulostulevan datan laskeminen, eli oppiminen



Autoencoder

- Autoencoder on malli, jossa sisään tuleva data muunnetaan pienempään dimensioon (encode). Vastaavasti data rekonstruktioidaan pienemmästä dimensiosta takaisin suurempaan (decode)
  - Autoencoder mahdollistaa epälineaariset kompleksiset muunnokset
- Rekonstruktion virhe pyritään saamaan mahdollisimman pieneksi, eli $X \approx \hat{X}$



Adversarial Autoencoder

- Makhzani, et. al, 2016, Adversarial Autoencoders, https://arxiv.org/pdf/1511.05644.pdf

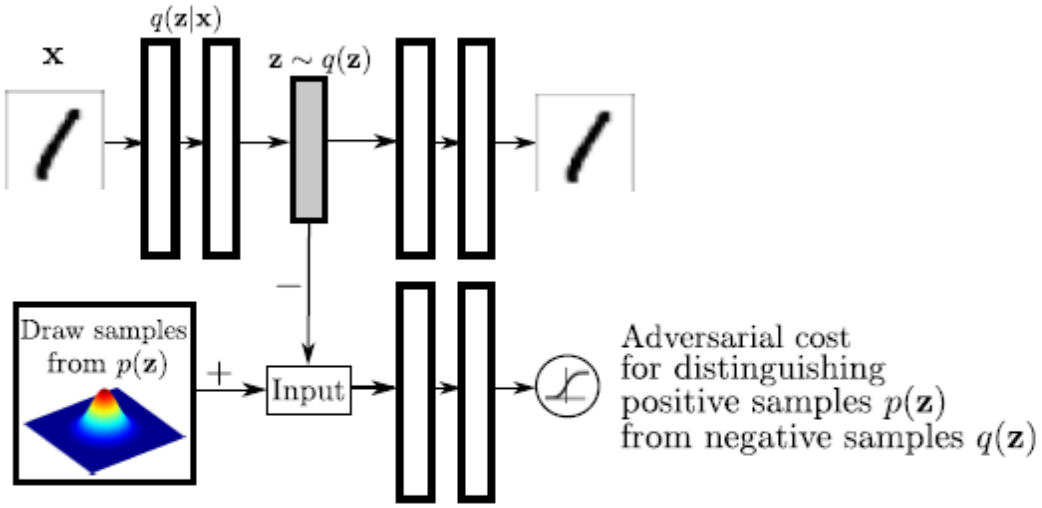
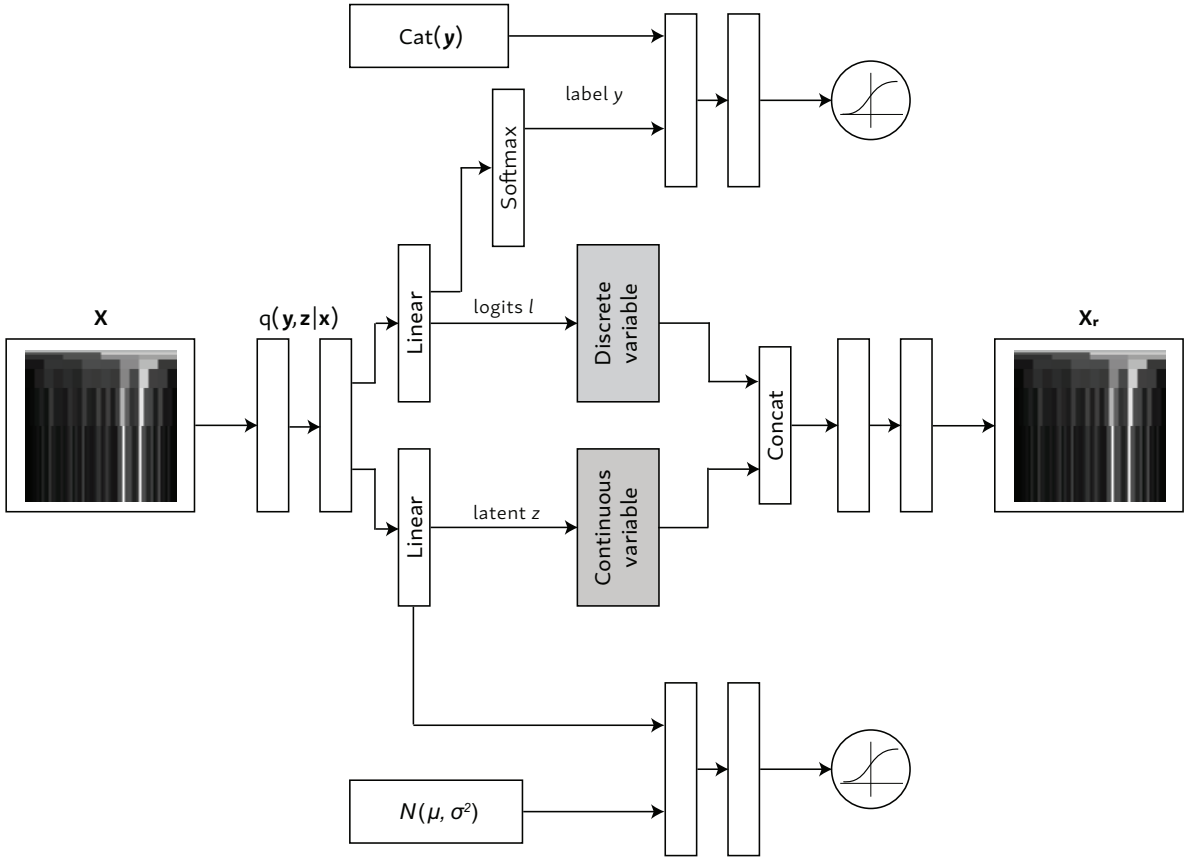


Figure 1: Architecture of an adversarial autoencoder. The top row is a standard autoencoder that reconstructs an image x from a latent code z. The bottom row diagrams a second network trained to discriminatively predict whether a sample arises from the hidden code of the autoencoder or from a sampled distribution specified by the user.

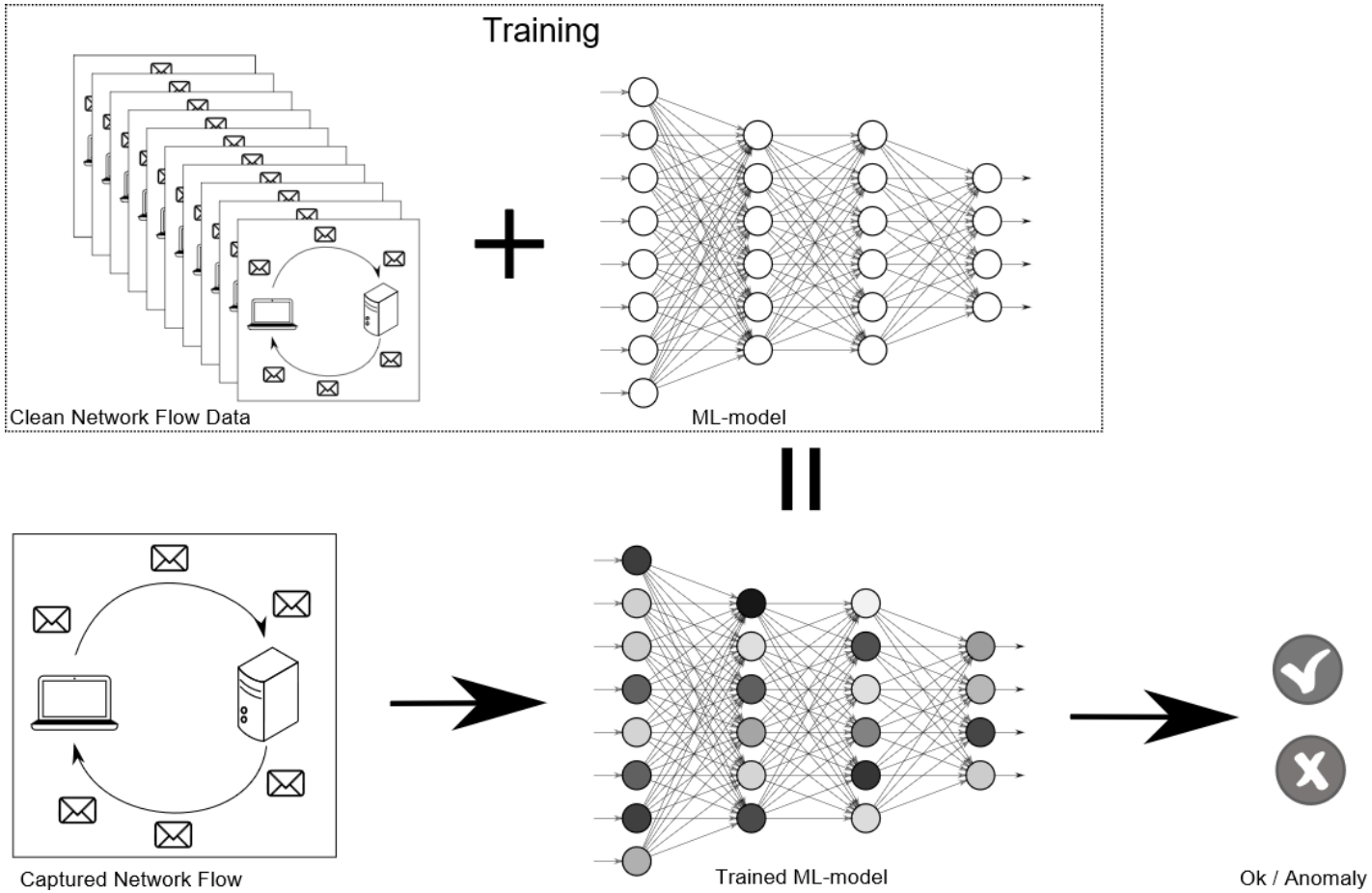
Adversarial Autoencoder, meidän malli



$$L(C, d_{\min}) = \frac{1}{|C|} \sum_{x \in C_i} \sum_{y \in C_j,\ i \neq j} \max\{0, d_{\min} - ||x - y||^2\}$$

Fig. 2: The general architecture of the neural network design used in this research. The encoder's outputs latent z and label y are enforced to match the selected distributions by establishing a two-player adversarial min-max game between the discriminator and the generator/encoder. [17]. The circles represent adversarial loss, while X_r represents the reconstruction.

Poikkeamien havainnointi

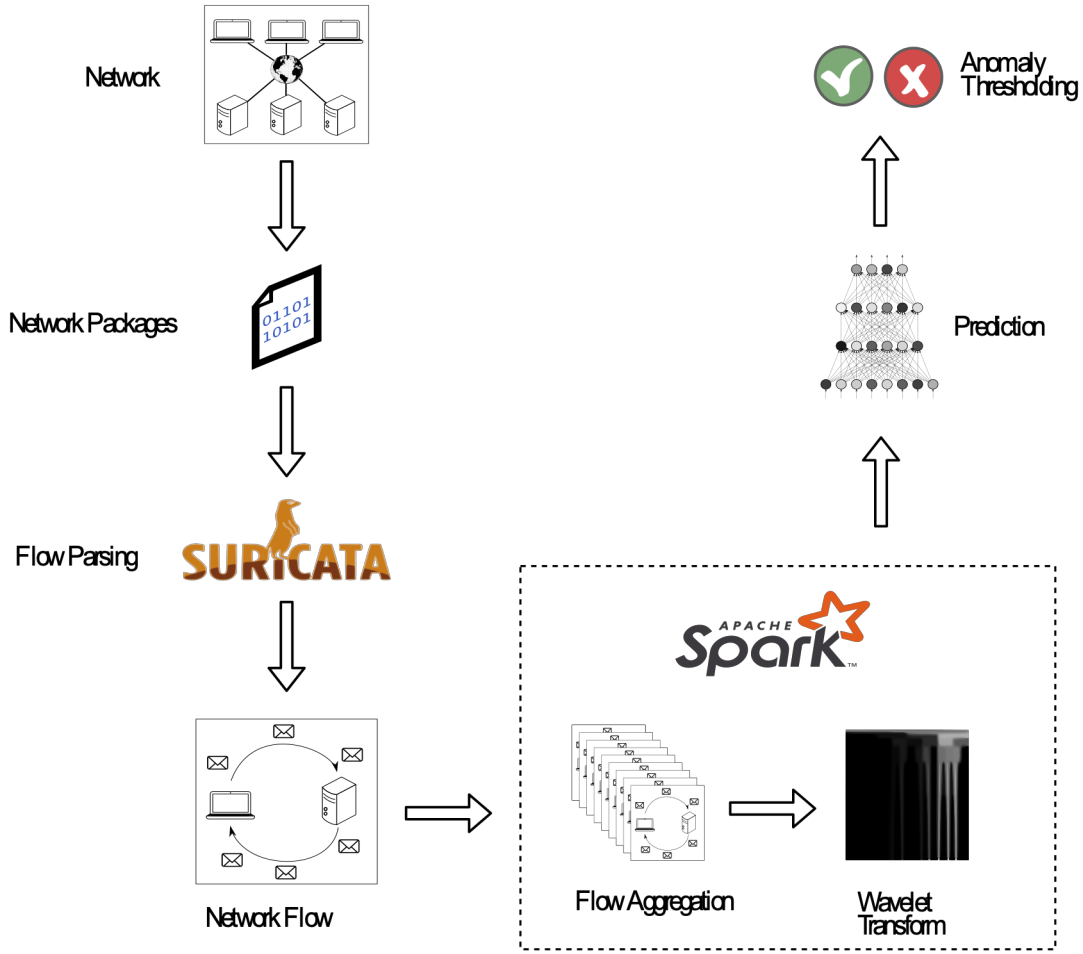


JYVSECTEC® Jyväskylä Security Technology

jamk.fi Institute of Information Technology

Sovelluksen toiminta

- Toteutuksessa käytettiin seuraavia OpenSource komponentteja
  - Suricata IDS/IPS/NMS -ohjelmistoa käytettiin pakettidatan parsimiseen
  - Apache Spark analyysimoottoria käytettiin ominaisuuksien valintaan (feature extraction) ja Wavelet muunnokseen
  - TensorFlow Machine Learning Framework:a käytettiin Adversarial Autoencoder osuuden toteutukseen

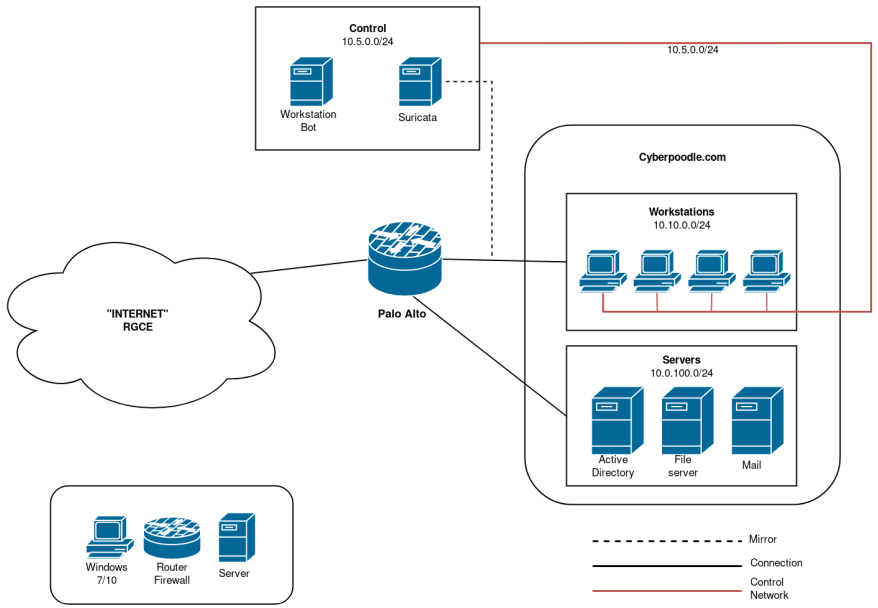


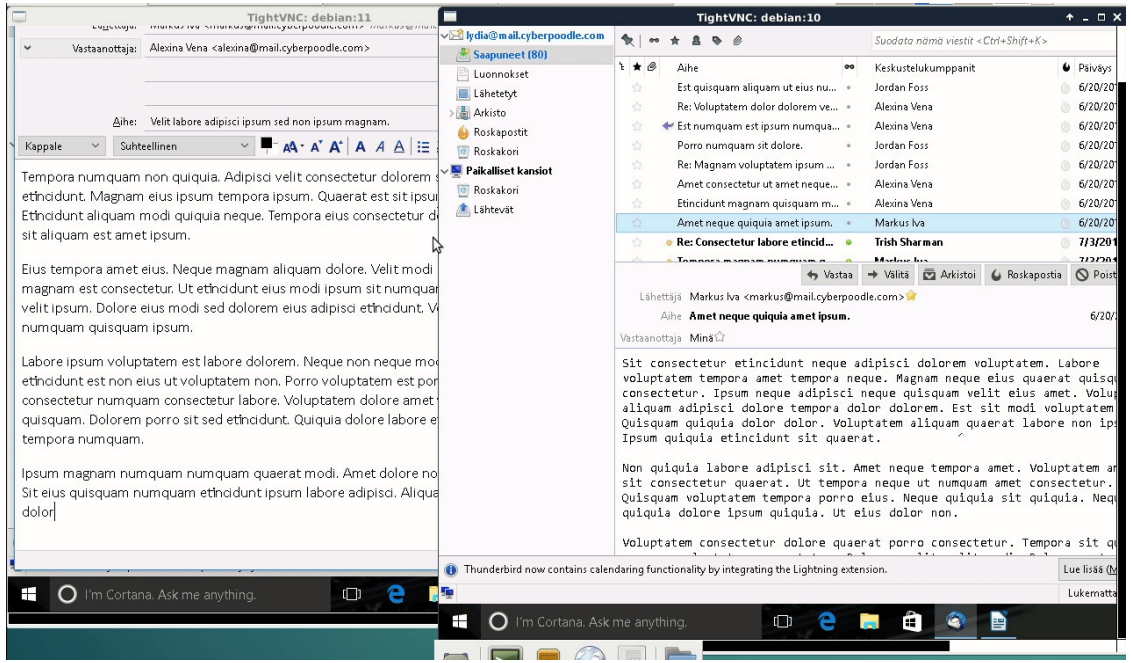
JYVSECTEC Jyväskylä Security Technology

jamk.fi Institute of Information Technology

Muuta kehitettyä

- Hankkeessa kehitettiin myös testiympäristö tarvittavan datan tuottamiseen. Botit käyttävät oikeita työpöytäsovelluksia.





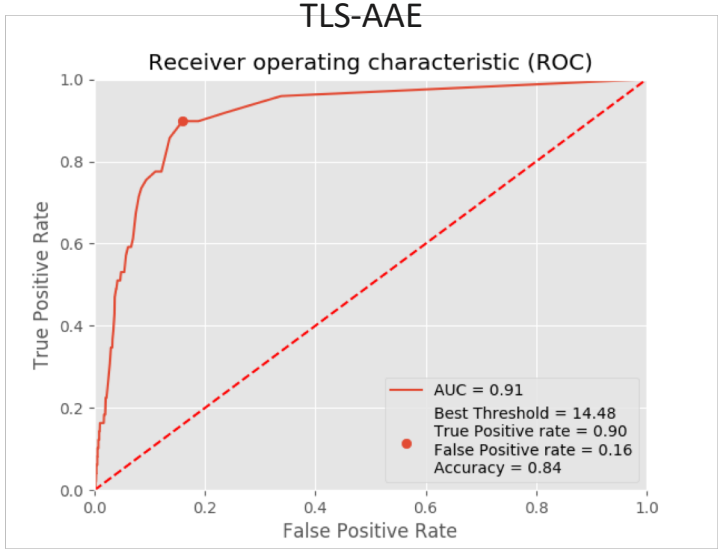
Havainnointituloksia

Adversarial Autoencoder (TLS-AAE)
vs
Convolutional Autoencoder (CAE)

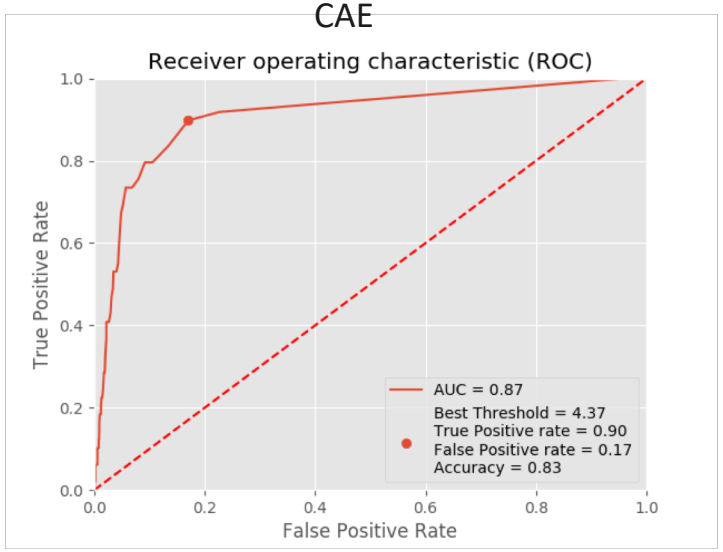
Performance characteristics

Method	TPR	FPR	Accuracy	AUC
TLS-AAE	0.90	0.16	84%	0.91
CAE	0.90	0.17	83%	0.87

TLS-AAE



CAE



JYVSECTEC Jyväskylä Security Technology

jamk.fi Institute of Information Technology

Yhteenveto ja jatkokehityskohteet

- Tulokset lupaavia, vaativat jatkokehitystä ja koneoppimismallien jatkokehitystä
- Testidata on yksi tärkeimmistä asioista koneoppimis- / tekoälykehityksessä
  - RGCE Cyber Range ja KYHA 18
- OpenSource sovelluksilla saadaan kehitettyä tekoälyyn perustuvaa poikkeamien tunnistukseen perustuvaa sensorimallia
- Jatkokehityskohteena mallin parantaminen ja tulosten visualisointi tukemaan päätöksentekoa

Tutkimuksen tuottamat tieteelliset julkaisut

- Puuska S., Kokkonen T., Alatalo J., Heilimo E., *"Anomaly-based Network Intrusion Detection using Wavelets and Adversarial Autoencoders"*, joka on esitelty 8-9.11.2019 International Conference on Information Technology and Communications Security SECITC-tapahtumassa (www.secitc.eu). Tutkimus julkaistaan Springerin Lecture Notes in Computer Science -kokonaisuudessa lähiaikoina (lopullisia ISBN-tietoja ei ole vielä saatavilla).
- Heilimo Eppu, Insinööri-opinnäytetyö, *"Applying Adversarial Autoencoders to Anomaly Detection in Cyber Security"*. Työ arvioidaan joulukuussa ja julkaistaan Internetissä Theseus-tietokannassa Jyväskylän ammattikorkeakoulun opinnäytetyönä.

Meidät tavoitat

www.jyvsectec.fi

www.jamk.fi

jyvsectec@jamk.fi

Seuraa meitä

Twitterissä @JYVSECTEC,

LinkedInissä ja YouTubessa

www.jyvsectec.fi/en

JYVSECTEC® Jyväskylä Security Technology

jamk.fi Institute of Information Technology