PENDETEKSIAN SERANGAN *IOT BOTNET* TIDAK DIKENAL DENGAN *UNSUPERVISED COMPETITIVE LEARNING*

TESIS

OLEH

SAID FAUZUL RUSNAIDI

NIM: 55419110011

PROGRAM MAGISTER TEKNIK ELEKTRO

FAKULTAS TEKNIK

UNIVERSITAS MERCU BUANA

2021

PENDETEKSIAN SERANGAN *IOT BOTNET* TIDAK DIKENAL DENGAN *UNSUPERVISED COMPETITIVE LEARNING*

TESIS

Diajukan sebagai Salah Satu Syarat untuk Menyelesaikan
Program Studi Magister Teknik

OLEH
SAID FAUZUL RUSNAIDI
NIM: 55419110011

PROGRAM MAGISTER TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS MERCU BUANA
2021

ABSTRAK

Pemanfaatan perangkat *IoT* sebagai *Robot Network* (*IoT Botnet*) merupakan ancaman serius pada *cybersecurity.* Hal tersebut tidak terlepas dari pertumbuhan perangkat *IoT* yang dilengkapi dengan teknologi komunikasi dan sistem komputasi yang canggih, namun tidak diiringi dengan kuatnya system keamanan yang diterapkan. Untuk mengantisipasi cepatnya pertumbuhan perangkat *IoT* dan/atau jenis *Botnet* di waktu yang akan datang, dibutuhkan penelitian untuk pendeteksian *IoT Botnet* pada perangkat *IoT* dan/atau jenis *Botnet* baru yang tidak dikenali sebelumnya (*Unknown Attack*). Dalam penelitian ini diusulkan pendeteksian *IoT Botnet* tidak dikenal menggunakan pendekatan *Unspervised Learning* yang diterapkan pada metode pembelajaran berbasis *Competitive Learning* (pada penelitian ini disebut dengan *Unsupervised Competitive Learning*). Pengujian dilakukan dengan metode *Learning Vector Quantization (LVQ)* dan memanfaatkan dataset *N-BaIoT: Data for network based detection of IoT botnet attack,* yang menyediakan data *real traffic* dari 9 perangkat *IoT* komersial yang telah diinfeksi *Bashlite/Gafgyt* dan *Mirai Botnet*. Pengujian dilakukan dalam 257 *batch* yang terbagi pada 5 skenario pengujian, yaitu pendeteksian *IoT Botnet* yang menggunakan data gabungan dari keseluruhan perangkat *IoT* dan *IoT Botnet* sebagai *baseline scenario* (Skenario-0) dengan nilai akurasi tertinggi didapatkan sebesar 99,54%, pendeteksian *IoT Botnet* sudah dikenal pada perangkat *IoT* sudah dikenal (Skenario-1) dengan rata-rata total akurasi didapatkan sebesar 85,22%, pendeteksian *IoT Botnet* sudah dikenal pada perangkat *IoT* belum dikenal (Skenario-2) dengan rata-rata total akurasi didapatkan sebesar 83,06%, pendeteksian *IoT Botnet* belum dikenal pada perangkat *IoT* sudah dikenal (Skenario-3) dengan rata-rata total akurasi didapatkan sebesar 79,98% dan pendeteksian *IoT Botnet* belum dikenal pada perangkat *IoT* belum dikenal (Skenario-4) dengan rata-rata total akurasi didapatkan sebesar 75,07%.

Kata Kunci: *IoT Botnet, Unknown Attack, Unsupervised Learning, Competitive Learning, Learning Vector Quantization.*

ABSTRACT

The use of IoT devices as a Robot Network (IoT Botnet) is a serious threat to cybersecurity. This is inseparable from the growth of IoT devices equipped with advanced communication technology and computing systems, but not accompanied by a strong security system in place. To anticipate the rapid growth of IoT devices and/or types of Botnets in the future, research is needed to detect IoT Botnets on new IoT devices and/or new types of Botnets (Unknown Attack) which have never been involved in the training process. In this study, We proposed to detect unknown IoT Botnets using the Unspervised Learning approach which is applied to Competitive Learning based method (hereinafter referred to as Unsupervised Competitive Learning). The test was carried out using Learning Vector Quantization (LVQ) method and the N-BaIoT dataset: Data for network based detection of IoT botnet attacks, which provides real traffic data from 9 commercial IoT devices that have been infected with Bashlite/Gafgyt and Mirai Botnet. Testing was carried out in 257 batches divided into 5 test scenarios, namely detection of IoT Botnet using combined data from all IoT devices and all kind of IoT Botnets as baseline scenario (Scenario-0) with the highest accuracy of 99.54%, detection of known IoT Botnets on known IoT devices (Scenario-1) with an average total accuracy of 85.22%, detection of known IoT Botnets on unknown IoT devices (Scenario-2) with an average total accuracy of 83.06%, detection of unknown IoT Botnet on known IoT devices (Scenario-3) with an average total accuracy of 79.98% and detection of unknown IoT Botnets on unknown IoT devices (Scenario-4) with an average total accuracy of 75.07%.

Keywords: *IoT Botnet, Unknown Attack, Unsupervised Learning, Competitive Learning, Learning Vector Quantization.*

PERNYATAAN *SIMILARITY CHECK*

Saya yang bertanda tangan di bawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh

Nama	: Said Fauzul Rusnaidi
NIM	: 55419110011
Program Studi	: Magister Teknik Elektro

dengan judul
"*Unknown IoT Botnet Attack Detection Using Unsupervised Competitive Learning*",
telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal 04/08/2021, didapatkan nilai persentase sebesar 18 %.

Jakarta, 04 Agustus 2021
Administrator Turnitin

Arie Pangudi, A.Md

PERNYATAAN

Saya yang bertandatangan di bawah ini menyatakan dengan sebenar-benarnya bahwa semua pernyataan dalam tesis ini :

Judul	:	Pendeteksian Serangan *IoT Botnet* Tidak Dikenal Dengan *Unsupervised Competitive Learning*
Bentuk Tesis	:	Penelitian Kuantitatif
Nama	:	Said Fauzul Rusnaidi
NIM	:	55419110011
Program	:	Magister Teknik Elektro
Tanggal	:	13 Agustus 2021

Merupakan hasil studi pustaka, penelitian lapangan, dan karya saya sendiri dengan bimbingan Dosen Pembimbing yang ditetapkan dengan Surat Keputusan Program Studi Magister Teknik Elektro Universitas Mercu Buana.

Karya ilmiah ini belum pernah diajukan untuk memperoleh gelar kesarjanaan pada program sejenis di perguruan tinggi lain. Semua informasi, data, dan hasil pengolahannya yang digunakan, telah dinyatakan secara jelas sumbernya dan dapat diperiksa kebenarannya.

Jakarta, 13 Agustus 2021

Said Fauzul Rusnaidi

PENGESAHAN

Judul	:	Pendeteksian Serangan *IoT Botnet* Tidak Dikenal Dengan *Unsupervised Competitive Learning*
Bentuk Tesis	:	Penelitian Kuantitatif
Nama	:	Said Fauzul Rusnaidi
NIM	:	55419110011
Program	:	Magister Teknik Elektro
Tanggal	:	14 Agustus 2021

Mengesahkan

Pembimbing

Dr. Marza Ihsan Marzuki, M.T.

Dekan Fakultas Teknik

Dr. Ir. Mawardi Amin, M.T

Ketua Program Studi
Magister Teknik Elektro

Dr. Umaisaroh, S.ST

KATA PENGANTAR

Alhamdulillah, puji syukur kehadirat Allah SWT yang senantiasa melimpahkan rahmat dan hidayah-Nya, sehingga tesis dengan judul "Pendeteksian Serangan *IoT Botnet* Tidak Dikenal Dengan *Unsupervised Competitive Learning*" ini dapat diselesaikan. Tesis ini disusun untuk memenuhi salah satu persyaratan menyelesaikan studi di Program Magister Teknik Elektro, Fakultas Teknik Universitas Mercu Buana.

Tentunya penyelesaian tesis ini tidak terlepas dari dukungan berbagai pihak, yang pada kesempatan ini penulis secara khusus menyampaikan rasa hormat dan terima kasih yang sebesar-besarnya, kepada:

1. Kedua Orang Tua, Istri dan Anak-anak yang senantiasa mendoakan dan memberikan dukungan baik moril maupun materil.
2. Bapak Dr. Marza Ihsan Marzuki, M.T., selaku dosen pembimbing yang telah menyediakan waktu, tenaga, dan pikiran dalam membimbing penulis.
3. Bapak Dr. Setiyo Budiyanto, S.T., M.Sc. dan Ibu Dr. Umaisaroh, S.ST. selaku penelaah dan penguji yang telah memberikan masukan dan saran membangun.
4. Bapak Prof. Dr. -Ing. Mudrik Alaydrus, Bapak Prof. Dr. Andi Andriansyah, M.Eng., serta semua dosen dan mahasiswa di Program Magister Teknik Elektro yang telah berbagi ilmu, pengalaman dan waktunya untuk berdiskusi.
5. Semua pihak yang telah membantu penyelesaian tesis ini, dengan tanpa mengurangi rasa hormat tidak dapat penulis sebutkan satu persatu.

Dengan keterbatasan ilmu dan pengalaman, penulis menyadari bahwa tesis ini masih banyak kekurangan dan membutuhkan pengembangan agar benar-benar dapat bermanfaat. Oleh sebab itu, penulis dengan senang hati menerima saran, kritik dan masukan yang membangun untuk kesempurnaan tesis ini.

Akhir kata, penulis berharap tesis ini dapat memberikan manfaat bagi kita semua terutama untuk pengembangan ilmu pengetahuan dan teknologi.

Jakarta, 14 Agustus 2021

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR NOTASI DAN ISTILAH

Perangkat *IoT* Tidak Dikenal	:	Perangkat *IoT* yang tidak digunakan dalam proses *training,* untuk mensimulasikan bahwa perangkat *IoT* tersebut saat ini belum tersedia di pasaran (baru akan hadir di waktu yang akan datang)
Perangkat *IoT* Sudah Dikenal	:	Perangkat *IoT* yang digunakan dalam proses *training,* untuk mensimulasikan bahwa perangkat *IoT* tersebut saat ini sudah tersedia di pasaran, sehingga dapat dipelajari pola/*patern* dari data trafiknya
IOT Botnet Tidak Dikenal	:	*IoT Botnet* yang tidak digunakan dalam proses *training,* untuk mensimulasikan bahwa *IoT Botnet* tersebut belum ada saat ini (baru akan hadir di waktu yang akan datang)
IOT Botnet Sudah Dikenal	:	*IoT Botnet* yang digunakan dalam proses *training,* untuk mensimulasikan bahwa *IoT Botnet* tersebut saat ini sudah ada dan diketahui keberadaannya (dalam hal ini *Mirai* dan *Gafgyt*), sehingga dapat dipelajari pola/*patern* dari data trafiknya
Basis pendeteksian	:	Perangkat *IoT* dan/atau *IoT Botnet* yang digunakan dalam proses *training*, dimana hasil *training* tersebut digunakan untuk pendeteksian *IoT Botnet* pada tahap *testing* sesuai dengan skenario yang ditetapkan
Obyek pendeteksian	:	Perangkat *IoT* dan/atau *IoT Botnet* yang dijadikan obyek pendeteksian *IoT Botnet* pada tahap *testing* sesuai dengan skenario yang ditetapkan
IoT	:	*Internet of Think*
Botnet	:	*Robot Network*

LVQ	:	*Learning Vector Quantization*
PRCL	:	*Pursuit Reinforcement Competitive Learning*
SVM	:	*Support Vector Machine*
LOF	:	*Local Outlier Factor*
NB	:	*Naïve Bayes*
KNN	:	*K-Nearest Neighbours*
RF	:	*Random Forest*
LR	:	*Logistic Regression*
DOS	:	*Denial of Service*
DDOS	:	*Distributed Denial of Service*
H	:	*Host Device*
MI	:	*MAC and IP address*
HH	:	*Source Host and Destination Host*
HH_jit	:	*HH Jitter*
HpHp	:	*Host and Port* dari *source* dan *destination*
$\mathbb{R}$	:	Anggota bilangan real
$StdDev$	:	Standar deviasi
x_i	:	*Data vector* x ke i
$x_1, \ldots, x_n$	:	*Data vector* x ke i sampai dengan *vector* x ke n
$y^1, \ldots, y^5$	:	*Subvector dari* x *yang ke 1 sampai subvector ke 5*
w_i	:	*Weight* atau bobot dari *neuron* yang ke i
$w_{i1}, \ldots, w_{id}$	:	*Weight* yang ke i untuk amatan *vector/subvector* ke 1 sampai ke d
m_i	:	Titik tengah dari *neuron (centroid)* yang ke i

$\|x - m_i\|$	:	Jarak antara *vector* x dan m yang ke i
c_p	:	*Class neuron* pemenang
c_r	:	*Class neuron runner up*
w_p	:	*Weight* neuron pemenang
w_r	:	*Weight* neuron *runner up*
$w^{(t-1)}$	:	*Weight* pada iterasi sebelumnya (*before*)
$w^{(t)}$	:	*Weight* pada iterasi terakhir (*current*)
$w^{(t+1)}$	:	*Weight* pada iterasi berikutnya (*next*)
α	:	*Parameter learning rate*, yaitu nilai yang menentukan berapa besar pengaruh hasil proses pendeteksian terhadap proses pendeteksian berikutnya (proses pembelajaran)
ε	:	*Parameter window*, yaitu nilai yang menentukan rentang area/daerah yang harus dipenuhi untuk memperbaharui *weight* dari *vector* pemenang dan *runner up*
$o_1, \ldots, o_n$	:	Output yang ke i sampai dengan output ke n
TP	:	*True Positive,* Hasil pendeteksian yang benar, dimana *IoT Botnet* terdeteksi sebagai *IoT Botnet*
TN	:	*True Negative,* Hasil pendeteksian yang benar, dimana data *benign* terdeteksi sebagai *benign*
FP	:	*False Positive,* Hasil pendeteksian yang salah, dimana data *benign* terdeteksi sebagai *Iot Botnet*
FN	:	*False Negative,* Hasil pendeteksian yang salah, dimana *IoT Botnet* terdeteksi sebagai data *benign*