ISSN : 2355-9365 e-Proceeding of Engineering : Vol.6, No.2 Agustus 2019 | Page 8808

Deteksi Serangan Denial of Service (DoS) menggunakan Algoritma Probabilistic Neural Network (PNN)

Tugas Akhir

diajukan untuk memenuhi salah satu syarat

memperoleh gelar sarjana

dari Program Studi Teknik Informatika

Fakultas Informatika Universitas

Telkom

1301154548

Astri Cahyaningtyas

Program Studi Sarjana Teknik Informatika

Fakultas Informatika

Universitas Telkom

Bandung

2019

LEMUAR PEN(;ESAIIAN

DctelL◊ Scnmi:an *Deniat of Service (l)oSJ* mcn!?1t11nruuon Algnrilma *I'rubabilislic Neural Nd"◊'"* (PNN I

Df!lectfon a/Ihnial of Serric« (Dos) Attacks u,inx tlte Probabilutil: Neural Netwarlc (PNN) Allt1rithm

NIM: 13'11154548

A◊trl CahyaninRtyas

Tugas aldur on, tclah dotcnma dan disahkan untuk mernenuhi sebagran syarai mcmpcrolch gelar pada Program Studt Sarjana Tcknik lnformatika

Fakultas lnforrnatika

Universitas Telkom

Bandung, 8 Apri l 2019

Mcnyctujui

Pcmbimb!p i.r /

Parmar! Sukarno. S.T.. M.Sc.. Ph.D.

NII': 17770073

Pembimbing 11

Muhammad Ancf Nugroho, S.T., M.T.

NIP: 15851886-2

Ketua Program Stmli

Sarjana Telmik lnformatika.

Nikcn Owi Wahyu Cahyani,S.T.. M.Kom.. Ph.D

NIP:00750052

LEMBAR PERNYATAAN

Dengan ini saya, Astri Cahyaningtyas, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul " Deteksi Serangan Denial of Service (DoS) menggunakan Algoritma Probabilistic Neural Network (PNN)" beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang belaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika dikemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya.

Bandung, 8 April 2019

Yang Menyatakan,

Astri Cahyaningtyas

Deteksi Serangan Denial of Service (DoS) menggunakan Algoritma Probabilistic Neural Network (PNN)

Astri Cahyaningtyas[1], Parman Sukarno[2], Muhammad Arief Nugroho[3]

[1,2,3] Fakultas Informatika, Universitas Telkom, Bandung
[1]astricahyaningtyas@students.telkomuniversity.ac.id, [2]psukarno@telkomuniversity.ac.id,
[3]arif.nugroho@telkomuniversity.ac.id

Abstrak

Salah satu masalah keamanan dalam jaringan komputer adalah serangan Denial of Service (DoS). Serangan DoS mengakibatkan pengguna dari akses layanan normal tidak dapat mengakses jaringan komputer dikarenakan penyerang mengonsumsi sumber daya yang berlebihan. Hal tersebut terjadi karena deteksi serangan DoS yang masih belum optimal. Untuk menyelesaikan masalah diatas, diusulkan membangun Intrusion Detection System (IDS) dengan metode anomaly-detection yang menggunakan algoritma machine learning yaitu Probabilistic Neural Network (PNN) untuk mendeteksi serangan DoS secara optimal. Pada penelitian ini, implementasi PNN dalam mendeteksi serangan DoS menggunakan NSL-KDD dataset dengan 13 fitur pilihan dan menghasilkan nilai akurasi tertinggi daripada algorima lain yaitu sebesar 98,06%.

Kata kunci : Keamanan Jaringan Komputer, Probabilistic Neural Network (PNN), Denial of Service (DoS).

Abstract

One of the security issue in computer network is Denial of Service (DoS) attack. DoS attack cause users who from normal service can not access network because attacker consumes excessive resources. It happens because the detection of DoS attacks is still not optimal. To solve the problem above, it is recommended to build Intrusion Detection System (IDS) with anomaly-detection that use machine learning algorithms, namely Probabilistic Neural Network (PNN) to improve DoS attacks optimally. In this study, the implementation of PNN for detecting DoS attacks using the NSL-KDD dataset with 13 selected features and produced the highest accuracy value than the other algorithms which is 98,06%.

Keywords: Computer Network Security, Probabilistic Neural Network (PNN), Denial of Service (DoS),

1. Pendahuluan

Latar Belakang

Meskipun ada banyak ancaman aktif di Internet, Denial of Serangan Service (DoS) adalah salah satu serangan paling umum di jaringan komputer. Serangan DoS mencegah pengguna dari akses layanan normal, dikarenakan penyerang mengonsumsi sumber daya jaringan komputer, memori, prosesor, dan lain-lain yang berlebihan [1]. Serangan DoS yang paling umum adalah ketika penyerang membanjiri(flood) jaringan komputer dengan banyak request pada saat bersamaan, membuat server tidak dapat merespon banyak request tersebut sehingga mengakibatkan pengguna yang sah tidak dapat berkomunikasi dengan server [1]. Intrusion Detection System (IDS) dan teknik pencegahan tradisional, seperti firewall, mekanisme kontrol akses, dan enkripsi, memiliki beberapa keterbatasan dalam melindungi jaringan dan sistem komputer dikarenakan serangan-serangan dalam jaringan komputer semakin canggih [12]. Selain itu, sebagian besar sistem yang dibangun berdasarkan teknik-teknik seperti itu memiliki tingkat deteksi positif palsu dan negatif palsu yang tinggi serta kurangnya adaptasi yang terus-menerus terhadap perubahan malicious behavior [12].

IDS adalah perangakat keamanan yang digunakan untuk memonitor lalu lintas jaringan komputer atau aktivitas sistem secara real-time dan akan mengirimkan peringatan kepada administrator atau mengambil beberapa tindakan aktif setelah serangan diidentifikasi [13]. Ada dua metode dalam melakukan intrusion detection yaitu signatured-based detection dan traditional anomaly-based detection. Kedua metode tersebut masih memiliki kelemahan diantaranya adalah low intelligence dan kemampuan adaptasi yang lemah pada skenario aplikasi yang berbeda-beda, serta kesulitan dalam memproses dataset yang berukuran sangat besar [13]. Dalam dekade terakhir, beberapa teknik machine learning telah diterapkan pada masalah intrusion detection dengan harapan dapat meningkatkan tingkat deteksi dan kemampuan beradaptasi [12]. Banyak teknik machine learning yang diperkenalkan dalam menangani masalah ini, yang paling banyak digunakan adalah Artificial Neural Network (ANN) [13]. Namun,

hasil penelitian menggunakan metode ANN-based masih tidak memuaskan dan tingkat ketelitian masih rendah [13].

Pada tugas akhir ini, penulis melakukan penelitian dalam mendeteksi serangan DoS menggunakan algoritma Probabilistic Neural Network (PNN). PNN merupakan suatu metode jaringan saraf tiruan (neural network) yang menggunakan pelatingan (training) supervised. Dalam penelitian ini, PNN digunakan untuk klasifikasi dengan dua kategori yaitu kategori normal dan kategori serangan DoS. Penulis memanfaatkan NSL-KDD sebagai dataset karena NSL-KDD dataset adalah dataset terbaik untuk menyimulasikan dan menguji performansi dalam pendeteksian serangan [4].

Topik dan Batasannya

Rumusan masalah dalam penelitian ini adalah karena metode IDS (signatured-based detection dan traditional anomaly-based detection) masih memiliki keterbatasan yaitu low intelligence dan tingkat adaptasi yang lemah, sehingga dibutuhkan metode IDS yang dapat mengalami masalah tersebut. Penelitian ini dilakukan untuk meningkatkan keamanan jaringan komputer dengan cara mendeteksi serangan DoS menggunakan algoritma Probabilistic Neural Network (PNN). Untuk menyimulasikan dan menguji performansi algoritma PNN, penulis menggunakan NSL-KDD dataset dengan 13 fitur terpilih.

Tujuan

Tujuan penelitian ini untuk meningkatkan keamanan jaringan komputer dengan cara mendeteksi serangan DoS menggunakan algoritma Probabilistic Neural Network (PNN). Keterkaitan antara tujuan, pengujian, dan kesimpulan dapat dilihat pada Tabel 1.

Tabel 1. Keterkaitan antara tujuan, pengujian dan kesimpulan

No	Tujuan	Pengujian	Kesimpulan
1	Meningkatkan keamanan jaringan komputer dengan cara mengimplementasikan algoritma PNN untuk deteksi serangan DoS dan menghasilkan akurasi.	Algoritma PNN berhasil mendeteksi serangan dengan menggunakan NSL-KDD dataset yang memiliki 113.270 data train dan 15.451 data test.	Hasil deteksi algoritma PNN memiliki akurasi sebesar 98,06%.

Organisasi Tulisan

Organisasi tulisan pada jurnal tugas akhir ini adalah sebagai berikut. Bab 1 mendeskripsikan pendahuluan yang berisi latar belakang, topik dan batasannya, tujuan, dan organisasi tulisan. Bab 2 membahas tentang studi terkait yang mendukung penulisan atau pengerjaan tugas akhir. Bab 3 memberikan penjelasan sistem yang dibangun. Hasil evaluasi yang menjadi tujuan dari penelitian ini ada di dalam Bab 4. Di akhir jurnal terdapat Bab 5 yang menarik kesimpulan dari penelitian ini.

2. Studi Terkait

Untuk memecahkan masalah network security diperlukan intrussion detection. Intrusion detection adalah suatu proses monitoring kejadian yang terjadi pada sistem atau jaringan komputer serta menganalisisnya untuk mengetahui aktivitas tersebut termasuk normal atau intrusi [5]. Metode intrusion detection diklasifikasikan menjadi tiga kategori: Signature-based Detection (SD), Anomaly-based Detection (AD) dan Stateful Protocol Analysis (SPA). Signature-based (SD) adalah pattern yang sesuai dengan serangan atau ancaman yang diketahui dengan cara membandingkan pattern terhadap peristiwa yang ditangkap untuk mengenali kemungkinan intrusi, signature-based juga dikenal dengan Knowledgebased Detection atau Misuse Detection karena menggunakan knowledge yang diakumulasikan oleh serangan dan vulnerability yang spesifik [5]. Anomaly-based detection (AD) atau yang disebut dengan Behavior-based detection adalah penyimpangan perilaku yang diketahui, dan profil mewakili normal atau ekspektasi perilaku yang berasal dari monitoring aktivitas reguler, koneksi jaringan komputer, host atau pengguna selama periode waktu tertentu [5]. Profil dapat berupa statis atau dinamis, dan dikembangkan untuk banyak atribut seperti upaya login yang gagal, penggunaan processor, jumlah email yang dikirim, dan lain-lain [5]. Kemudian, AD membandingkan profil normal dengan peristiwa yang diamati untuk mengenali serangan signifikan[5]. Anomaly-detection mengalami peningkatan popularitas karena menjadi efektif terhadap serangan baru yaitu dengan memanfaatkan algortima machine learning [2]. Ada banyak algoritma klasifikasi dalam machine learning yang dilatih dan digunakan untuk deteksi serangan dalam jaringan komputer, untuk lebih meningkatkan kinerja

pengklasifikasian dan mengurangi waktu deteksi, dapat dilakukan dengan cara mengurangi fitur-fitur [2]. Stateful di SPA menunjukkan bahwa IDS dapat mengetahui dan melacak status protokol dimana SPA adalah profil generin yang dikembangkan untuk protokol tertentu [5].

Berdasarkan penelitian [13], Probabilistic Neural Network (PNN) dapat digunakan untuk mendeteksi serangan yang terdapat dalam jaringan komputer. PNN hanya memerlukan satu forward process dan tidak mempunyai back propagation. Hasil eksperimen dengan menggunakan KDDCUP99 dataset menunjukan metode PNN memiliki rata-rata performansi yang lebih baik daripada Decision Tree, Naive Bayes, dan BPNN. Untuk serangan DoS, metode PNN menghasilkan nilai Precision 99,9%, nilai Recall adalah 99,79%, dan F-value sebesar 99,84%. Nilai-nilai tersebut lebih besar daripada metode yang lain.

Pada penelitian [3], dengan menggunakan MATLAB membuktikan bahwa PNN memberikan nilai akurasi yang lebih baik daripada Feed Forward Neural Network (FFNN) and Radial Basis Neural Network (RBNN). Dengan mengurangi 41 fitur menjadi 13 fitur berdasarkan Principal Component Analysis (PCA), nilai akurasi PNN meningkat menjadi 97,5%.

Penelitian [7] menjelaskan bahwa kategori klasifikasi juga berpengaruh pada nilai akurasi dari algoritma PNN dalam mendeteksi serangan. Nilai akurasi dengan lima kategori adalah 97,5% sedangkan terjadi peningkatan nilai akurasi jika hanya menggunakan dua kategori yaitu sebesar 98%.

NSL-KDD merupakan perbaikan dari KDDCUP99 yang banyak digunakan dalam penelitian intrusion detection sekaligus sebagai validasi dari algoritma yang digunakan. KDDCUP99 diperbaiki karena terdapat kekurangan yaitu adanya record yang redundannya besar (yang dapat mengakibatkan algoritma machine learning menjadi bias) dan kompleksitasnya yang tinggi. Dari hasil penelitian [8], metode Ensemble-based multi-filter feature selection (EMFFS) adalah yang terbaik dalam melakukan preprocessing dataset untuk memilih fitur-fitur penting dalam pengklasifikasian dari suatu algoritma. Dibandingkan dengan metode lainnya, EMFFS menghasilkan performansi terbaik dengan nilai akurasi 99,67%, detection rate 99,76%, false alarm rate 0,42%, dan time 0,78s.

Serangan DoS dimaksudkan untuk menghentikan service target yang disediakan dengan flooding illegitimate requests. Oleh karena itu, agar serangan DoS terdeteksi, traffic feature seperti ”persentase koneksi memiliki host tujuan yang sama dan layanan yang sama ” dan packet level feature seperti ”source byte” dan ”persentase paket error” adalah fitur-fitur yang sangat signifikan.

Dalam serangan DoS seperti flooding diperlukan penelurusan terhadap traffic yang masuk ke dalam jaringan komputer. Fitur service dimanfaatkan untuk mencatat service/layanan apa yang digunakan di jaringan tujuan, contohnya seperti ftp_data, http, telnet, dan lain-lain. Fitur flag dapat membantu dalam penentuan apakah traffic tersebut termasuk serangan DoS atau bukan karena jika traffic tersebut memiliki flag yang sama dalam jumlah banyak dalam frekuensi waktu tertentu dapat dipastikan, penyerang sedang melakukan flooding, contoh flagnya adalah SF yang artinya koneksi normal (SYN/FIN selesai) berbeda dengan REJ yang berarti koneksi ditolak (SYN terlihat tetapi balasannya adalah RST). src_bytes dan dst_bytes untuk menyimpan jumlah data byte yang dikirim, karena jika jumlah data byte besar maka traffic tersebut dapat dicurigai mengandung serangan DoS. Status logged_in diperlukan untuk mengetahui apakah berhasil log in atau tidak, fitur ini tidak terlalu signifikan tetapi dapat membantu dalam menentukan apakah termasuk serangan DoS atau bukan. Fitur count akan menghitung berapa banyak sumber tersebut melakukan koneksi ke host tujuan yang sama dalam waktu dua detik terakhir, semakin besar nilai count semakin besar pula kemungkinan data tersebut diklasifikasikan sebagai serangan DoS. Presentase koneksi yang telah mengaktifkan flag atau fitur serror_rate dapat mendukung data traffic tersebut dalam menentukan label serangan DoS atau bukan. Fitur same_srv_rate dan dst_host_same_srv_rate adalah fitur yang paling signifikan dalam penentuan label serangan DoS atau bukan karena fitur ini mencatat persentase koneksi dengan service yang sama dimana traffic tersebut dikatakan sedang terjadi flooding karena ada sumber yang sedang melakukan koneksi dengan service sama. Sehingga semakin besar nilai same_srv_rate dan dst_host_same_srv_rate-nya maka semakin besar pula kemungkinan data tersebut adalah serangan DoS. Kebalikan dari same_srv_rate, diff_srv_rate untuk menentukan persentasi koneksi dengan service yang berbeda. dst_host_count untuk mencatat jumlah koneksi ke host tujuan yang memiliki IP address yang sama sedangan untuk mencatat port number ada pada fitur dst_host_srv_count. Fitur dst_host_serror_rate dan dst_host_srv_serror_rate akan mendukung data yang lain dalam menentukan serangan DoS atau bukan. 13 fitur tersebut ada yang signifikan ada yang tidak, sehingga 13 fitur ini dibutuhkan untuk saling mendukung dan saling melengkapi dalam menentukan apakah traffic tersebut merupakan serangan DoS atau bukan.

Seperti yang ditunjukkan pada Gambar 1 dan Algoritma 1, metode PNN adalah feed forward neural network dengan 4 layer adalah Input Layer, Hidden Layer, Summation Layer, dan Output Layer. Input Layer menerima inputan data test. Input Layer adalah vektor dengan banyak feature. Sehingga jumlah neuron di Input Layer sama dengan dimensi vektor input. Neuron-neuron di Input Layer akan dihitung menggunakan fungsi 1 untuk menjadi node-node di Hidden Layer. Dengan memfokuskan pada Gaussian yang digunakan untuk mewakili fungsi kepekatan probabilitas variabel acak distribusi normal, perhitungan yang diperlukan untuk inferensi dan pembelajaran menjadi relatif mudah sehingga masalah pembelajaran yang diawasi dalam pembelajaran mesin yang dapat

Tabel 2. Dataset Features

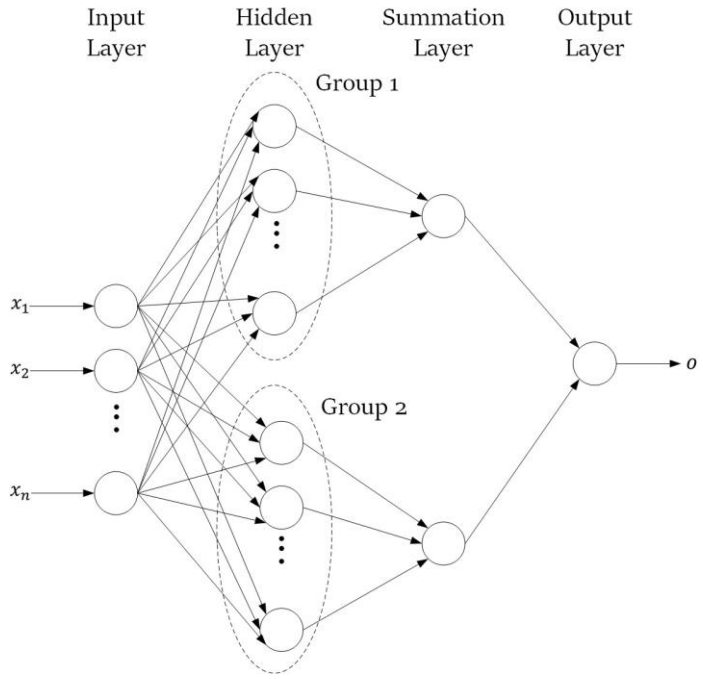
Atribut	Deskripsi [4]
service	Service jaringan tujuan yang digunakan.
flag	Status koneksi - normal atau eror.
src bytes	Jumlah data byte yang dikirim dari sumber ke tujuan dalam single connection.
dst bytes	Jumlah data byte yang dikirim dari tujuan ke sumber dalam single connection.
logged in	Status logged in: 1 jika sukses, 0 sebaliknya.
count	Jumlah koneksi ke host tujuan yang sama dengan koneksi saat ini dalam dua detik terakhir.
serror rate	Persentase koneksi yang telah mengaktifkan flag s0, s1, s2 atau s3, diantara koneksi-koneksi yang telah terhitung.
same srv rate	Persentase koneksi dengan service yang sama, diantara koneksi-koneksi yang telah terhitung.
diff srv rate	Persentase koneksi dengan service yang berbeda, diantara koneksi-koneksi yang telah terhitung.
dst host count	Jumlah koneksi yang IP address tujuan yang sama.
dst host srv count	Jumlah koneksi yang memiliki port number yang sama.
dst host same srv rate	Persentase koneksi yang memiliki service yang sama, diantara koneksi-koneksi yang terintegrasi dengan dst host count.
dst host serror rate	Persentase koneksi yang telah mengaktifkan flag s0, s1, s2 atau s3, diantara koneksi-koneksi yang terintegrasi dengan dst host count.
dst host srv serror rate	Persentase koneksi yang telah mengaktifkan flag s0, s1, s2 atau s3, diantara koneksi-koneksi yang terintegrasi dengan dst host srv count.

dianggap sebagai pembelajaran fungsi dari contoh dapat dilemparkan langsung ke dalam kerangka Gaussian [9]. Parameter yang berpengaruh dalam perhitungan fungsi Gaussian adalah standar deviasi (std), jika nilai std kurang tepat maka pesebaran data menjadi tidak bagus yang mengakibatkan hasil output akhir dari pengklasifikasian menjadi salah. Sehingga diperlukan ekperimen untuk mengetahui berapa nilai std yang tepat bagi pattern yang telah ada. Hidden Layer adalah bagian terpenting dalam network. Neuron di Hidden Layer didapatkan dari hasil perhitungan menggunakan fungsi Gaussian, kemudian dibagi menjadi dua kelas yaitu group 1 dan group 2. Group 1 digunakan untuk memproses input pattern yang normal dan group 2 digunakan untuk memproses input pattern yang abnormal. Jumlah neuron di Hidden Layer sama dengan jumlah training pattern. Summation Layer mempunyai 2 neuron, satu untuk menunjukkan operasi penjumlahan dari group 1 dan yang lain untuk neuron group 2. Summation Layer merupakan penjumlahan untuk setiap kelas menggunakan fungsi 2. Output Layer hanya mempunyai satu neuron. Setelah mengumpulkan dua output dari Summation Layer kemudian dilakukan perbandingan. Pada akhirnya, testing pattern menemukan kelas klasifikasinya berdasarkan prinsip ”bigger is winner”.

Confusion matrix[11] adalah tabel yang sering digunakan untuk menggambarkan kinerja model deteksi pada suatu data train. Matriks ini berisi nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Rumus 3 digunakan untuk menghitung akurasi sebagai nilai seberapa bagus kinerja model deteksi menggunakan algoritma PNN. Semakin tinggi nilai akurasi berarti algoritma PNN adalah algoritma yang bagus untuk mendeteksi serangan DoS. Ini dapat memudahkan administrator jaringan komputer untuk mengetahui bahwa jaringan tersebut sedang ada serangan atau tidak.

$$\frac{TP + TN}{TP + TN + FP + FN} * 100\% \quad (3)$$

- True Positive (TP) adalah jumlah catatan serangan yang dideteksi sebagai serangan. Misalnya, hasil klasifikasi realitasnya adalah DoS dan hasil prediksi model machine learning adalah DoS, maka nilai TP bertambah 1.
- True Negative (TN) adalah jumlah catatan normal yang dideteksi sebagai normal. Misalnya, hasil klasifikasi realitasnya adalah normal dan hasil prediksi model machine learning adalah normal, maka nilai TN bertambah 1.
- False Positive (FP) adalah jumlah catatan normal yang dideteksi sebagai serangan. Misalnya, hasil klasifikasi



Gambar 1. Arsitektur PNN

Algorithm 1 Probabilistic Neural Network (PNN)

1: Membaca data train dan data test
2: Membaca neuron input . Input Layer
3: Hitung menggunakan fungsi Gaussian . Hidden Layer

$$\exp^{\frac{-((x-x_0)^2+(y-y_0)^2)}{2\sigma^2}} \quad (1)$$

4: Menghitung jumlah dari fungsi Gaussian dari setiap kategori . Summation Layer

$$\sum_{i=1}^{n} \exp^{\frac{-((x-x_i)^2+(y-y_i)^2)}{2\sigma^2}} \quad (2)$$

5: Memilih nilai terbesar dari neuron-neuron yang ada di Summation Layer . Output Layer

realitasnya adalah normal dan hasil prediksi model machine learning adalah DoS, maka nilai FP bertambah 1.

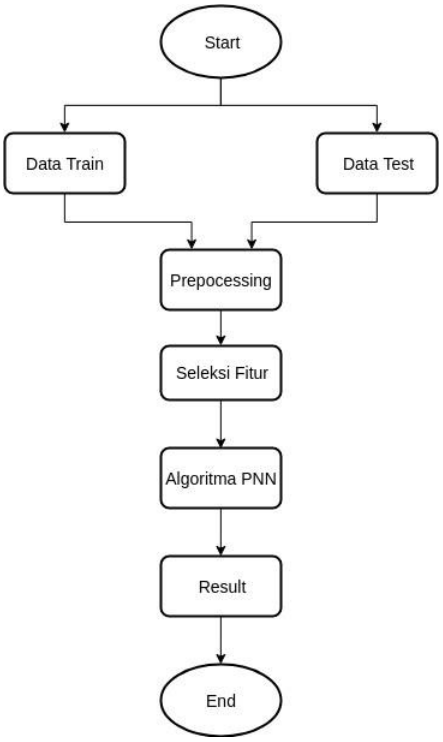
- False Negative (FN) adalah jumlah serangan yang sebenarnya namun dideteksi sebagai normal. Misalnya, hasil klasifikasi realitasnya adalah DoS dan hasil prediksi model machine learning adalah normal, maka nilai FN bertambah 1.

Pada kasus ini nilai yang penting untuk security[6] adalah nilai False Negative. Nilai ini berhubungan dengan kesalahan algoritma dalam melakukan deteksi serangan. Semakin kecil nilai False Negative semakin bagus algoritma tersebut dalam deteksi serangan DoS.

3. Sistem yang Dibangun

Pada penelitian ini digunakan dataset NSL-KDD yang berupa data train dan data test dengan split data 20% untuk data test dan 80% untuk data train. Pada dataset ini terdiri 41 fitur dengan 4 kategori serangan[4]. Pada penelitian ini, memakai 15.451 data test dan 113.270 data train dan hanya berfokus pada jenis serangan DoS.

Preprocessing dilakukan untuk memudahkan algoritma dalam membaca dataset. Pada tahap ini, jenis serangan akan di kategorikan sebagai DoS dan normal, dimana serangan DoS adalah back, land, neptune, pod, smurf, dan



Gambar 2. Flow Chart Sistem

teardrop, kemudian mengubah semua atribut kedalam format numeric. Contoh: TCP diubah menjadi 1, HTTP diubah menjadi 2, UDP diubah menjadi 3, dan sebagainya.

Tahap ini dilakukan seleksi fitur agar mendapatkan hasil yang optimal. Dari proses seleksi fitur[8], terdapat 13 fitur yang paling relevan terhadap serangan DoS (dapat dilihat pada Tabel 2).

Tahap selanjutnya adalah deteksi serangan menggunakan algoritma PNN yang mengoutputkan nilai-nilai TP, TN, FP, dan FN dan menghasilkan nilai akurasi sesuai dengan rumus 3. Jika nilai TP, TN, FP, dan FN ditambahkan maka hasilnya adalah banyaknya data train.

4. Evaluasi

4.1 Hasil Pengujian

Hasil pengujian algoritma PNN untuk mendeteksi serangan dengan memanfaatkan NSL-KDD dataset dapat dilihat pada Tabel 3. Sebagai perbandingan hasil deteksi serangan pada jaringan komputer, ditambahkan algoritma Naive Bayes, Decision Tree, dan model probabilitas selain Gaussian yaitu model Stokastik. Alasan algoritma Naive Bayes dan PNN dibandingkan adalah mereka sama-sama menggunakan teori Bayesian dalam menghitung setiap probabilitas. Sedangkan algoritma Decision Tree dan PNN dibandingkan karena Decision Tree tidak butuh melakukan perubahan data numerical karena Decision Tree tidak memerlukan perhitungan matematis, kebalikan dari PNN yang memerlukan perhitungan matematis. Kemudian dilakukan juga eksperimen terhadap model probabilitas Stokastik Gradient Descent (SGD) untuk mengetahui apakah model probabilitas Gaussian lebih baik atau tidak. SGD merupakan pendekatan classifier yang cukup sederhana dan dapat diaplikasikan pada large-scale dan sering digunakan untuk klasifikasi teks dan Natural Language Processing (NLP).

Tabel 3. Hasil pengujian algoritma PNN

Algoritma	Akurasi	TP	TN	FP	FN
Naive Bayes	90,38%	4.831	9.134	576	910
Decision Tree	96,11%	5.654	9.196	514	87
PNN	98,06%	5.584	9.567	143	157
SGD	90,88%	4.932	9.110	600	809

4.2 Analisis Hasil Pengujian

Input layer memiliki 13 neuron karena setiap record yang ada di dalam dataset memiliki 13 fitur. Hidden Layer memiliki 113.270 neuron, sebanyak record yang ada di data train. PNN merupakan proses one feed forward dan tidak ada back propagation. Summation layer memiliki 2 neuron yang merepresantikan jumlah kategori pengklasifian dan output layer memiliki 1 neuron yang merupakan hasil akhir keputusan. Dapat dilihat pada Tabel 3 algoritma PNN memiliki akurasi tertinggi yaitu 98,06%. Nilai TP (True Positive) merupakan catatan serangan yang dideteksi sebagai serangan, hasil pengujian menghasilkan 5.584 data yang teridentifikasi sebagai TP. Sebanyak 9.567 teridentifikasi sebagai TN (True Negative), TN adalah jumlah catatan normal yang terdeteksi normal. Nilai FP (False Positive) sebanyak 600 yang artinya ada catatan normal yang dideteksi sebagai serangan. Nilai FN (False Negative) adalah jumlah serangan yang sebenarnya namum dideteksi sebagai normal dan di pengujian ini, nilai FN sebanyak 809 data. Algoritma PNN menghasilkan nilai akurasi yang lebih baik daripada yang lain karena perhitungan PNN sangat detail dan berbeda daripada algortima yang lain, terlihat pada semua neuron yang ada di input layer dihitung dengan model Gaussian terhadap semua neuron yang ada di data train. Hal tersebut dapat membantu meminimalisir kesalahan dalam proses pengklasifikasian. Model Gaussian juga sangat mendukung algoritma PNN ini, karena terbukti saat dibandingkan dengan SGD, nilai akurasi Gaussian jauh lebih baik. Perhitungan Gaussian bergantung pada nilai standar deviasi telah ditentukan berdasarkan hasil beberapa ekperimen, pada penelitian ini nilai std=0,001. Nilai ini sangat berpengaruh karena mempengaruhi pesebaran data. Fungsi Gaussian dengan distribusi normal dinilai lebih optimal dibandingkan SGD.

Nilai akurasi PNN merupakan nilai yang tertinggi daripada algortima lainnya dan PNN memiliki kelebihan yang tidak dimiliki oleh algortima lain. Kelebihan PNN [10] adalah stuktur paralel yang inheren (berhubungan erat), dapat dilihat dalam proses algoritma PNN dimana semua atribut data input akan dihitung secara matematis dengan seluruh atribut yang ada di data train sehingga dapat dipastikan bahwa record data test akan diuji serinci mungkin pada data train untuk meminimalisir kesalahan dalam proses pengklasifikasian. PNN juga menjamin menjadi classifer yang optimal karena ukuran data train yang terus meningkat. Sampel training pada PNN dapat ditambahkan atau dihapus tanpa pelatihan ulang yang intensif.

5. Kesimpulan

Jaringan komputer saat ini harus diamankan dari serangan-serangan yang semakin canggih. Intrusion Detection System (IDS) memegang peranan penting dalam menyelesaikan masalah keamanan jaringan komputer. Penelitian ini membuktikan bahwa machine learning dapat meningkatkan performansi IDS jika dibandingkan dengan metode tradisional. Beberapa model algoritma telah diuji untuk menentukan akurasi yang terbaik. Dan nilai akurasi yang paling unggul adalah algoritma Probabilistic Neural Network (PNN) yaitu 98,06%. Sehingga dapat disimpulkan bahwa PNN lebih baik dalam mengatasi masalah keamanan jaringan komputer dan dapat meningkatkan keamanan jaringan komputer.

Saran untuk penelitian selanjutnya adalah melakukan eksperimen terhadap data real traffic untuk mengevaluasi kembali apakah algoritma PNN sesuai dengan environment sistem jaringan komputer tersebut.

Daftar Pustaka

[1] A review of machine learning techniques efficiency in dos attack detection. International Journal of Scientific Research, 6:461–462, 2017.

[2] S. K. Biswa. Intrusion detection using machine learning: A comparison study. International Journal of Pure and Applied Mathematics, 119:101–114, 2018.

[3] S. Devaraju and D. S. Ramakrishnan. Performance analysis of intrusion detection system using various neural network classifier. International Conference on Recent Trends in Information Technology (ICRTIT), pages 1033–1038, 2011.

[4] L. Dhanabal and P. Shantharajah. A study on nsl-kdd dataset for intrusion detection system based on classification algorithms. International Journal of Advanced in Computer and Communication Engineering, 4:446–452, 2015.

[5] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung. Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications, 36:16–24, 2013.

[6] A. Milenkoski, M. Vieira, S. Kounev, A. Avritzer, and B. D. Payne. Evaluating computer intrusion detection systems: A survey of common practices. ACM Computing Surveys, 48:12:1–12:41, 2015.

[7] L. Ning. Network intrusion classification based on probabilistic neural network. International Conference on Computational and Information Sciences, pages 57–59, 2013.

[8] O. Osanaiye, H. Cai, K.-K. R. Choo, A. Dehghantanha, Z. Xu, , and M. Dlodlo. Ensemble-based multi-filter feature selection method for ddos detection in cloud computing. EURASIP Journal on Wireless Communications and Networking, pages 1–10, 2016.

[9] C. E. Rasmussen and C. K. I. Williams. Gaussian processes for machine learning. The MIT Press, pages 1–266, 2006.

[10] S. S. Sawant and P. S. Topannavar. Introduction to probabilistic neural network. International Journal of Advanced Research in Computer Science and Software Engineering, 5:279–283, 2015.

[11] D. School. Simple guide to confusion matrix terminology. http://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/, 2008. Online; Accessed 5 April 2019.

[12] M. Zamani. Machine learning techniques for intrusion detection. arXiv Computer Science, pages 1–10, 2013.

[13] M. Zhang, J. Guo, B. Xu, and J. Gong. Detecting network intrusion using probabilistic neural network. 11th International Conference on Natural Computation (ICNC), pages 1151–1158, 2016.