ANALISIS *INTRUSION DETECTION SYSTEM* DI INTERNAL JARINGAN WAN MENGGUNAKAN *DATA MINING*: STUDI KASUS PADA ASTRIDO GROUP JAKARTA

INTRUSION DETECTION SYSTEM ANALYSIS IN INTERNAL WAN NETWORK USING DATA MINING: A CASE STUDY IN JAKARTA ASTRIDO GROUP

Agni Isador Harsapranata

AMIK Bina Sarana Informatika Bekasi
agniisador@gmail.com

Abstrak

Keamanan informasi merupakan salah satu bagian penting dalam suatu sistem informasi, karena informasi merupakan aset yang sangat berharga bagi setiap institusi ataupun perusahaan. Informasi sangat penting sehingga informasi hanya diperuntukkan bagi orang-orang yang memiliki wewenang saja. Oleh karena itu, keamanan sistem informasi harus terjamin sehingga orang yang tidak berkepentingan tidak akan bisa mendapatkan akses terhadap informasi tersebut. Pengamanan informasi dari pihak-pihak yang tidak berkepentingan saat ini sudah berkembang dengan pesat, salah satunya menggunakan *Intrusion Detection System* (IDS). Dengan menggunakan peralatan tersebut setiap kali ada percobaan *intrusion* oleh orang yang tidak memiliki otoritas, akan ada peringatan yang dikirimkan ke pihak administrator. Tindakan deteksi intrusi adalah upaya untuk memantau dan dimungkinkan untuk mencegah berbagai upaya menyusup ke sistem dan sumber daya informasi perusahaan. Deteksi intrusi adalah untuk menemukan pihak yang tidak memiliki otoritas ke sumber daya informasi, dengan cara melakukan pengamatan di jaringan perusahaan, dikarenakan saat ini pertumbuhan jaringan komputer yang besar, semakin mudahnya mendapatkan berbagai aplikasi dan alat untuk melakukan kegiatan intrusi, dengan demikian deteksi terhadap intrusi menjadi suatu hal yang sangat penting. Dalam penelitian ini akan digunakan metode *support vector machine*. Setelah dilakukan pengujian maka hasil yang didapat adalah *support vector machine* menghasilkan nilai akurasi sebesar 86,30 %, nilai *Sensitivity* 83,80% dan nilai AUC sebesar 0,857 untuk 10 menit pertama, sebesar 76,70 %, nilai *Sensitivity* 77,51% dan nilai AUC sebesar 0,858 untuk 10 menit kedua, sebesar 82,94 %, nilai *Sensitivity* 79,75% dan nilai AUC sebesar 0,818 untuk 10 menit ketiga .

Kata kunci: *intrusion detection system, WAN, data mining.*

Abstract

Information security is one important part in information system, because information is a valuable asset to any institution or company. Information is very important so that it is directed only to authorized persons. Therefore, information systems security should be guarded so that unauthorized persons will not be able to access this information. Securing information from unauthorized access is now growing rapidly, e.g. using Intrusion Detection System (IDS). With this equipment, each time there is an attempted intrusion by unauthorized persons, there will be a warning sent to the administrator. Intrusion detection is to monitor and prevent various attempts to infiltrate the system and corporate information resources. Intrusion detection is to find unauthorized persons who try to access information resources by observing the corporate network. With the growth of large computer network, it is easier to acquire applications and tools to perform intrusion activities, thus intrusion detection gains importance. In this study, support vector machine method will be used. The results show that support vector machine produces an

accuracy rate of 86.30%, sensitivity value of 83.80%, and AUC value of 0.857 for the first 10 minutes, amounting to 76.70%; sensitivity value of 77.51%, and AUC value of 0.858 for the second 10 minutes, amounting to 82.94%; sensitivity value of 79.75% and AUC value of 0.818 for the third 10 minutes.

Keywords: *intrusion detection system,* WAN, *data mining*

Tanggal Terima Naskah : 29 Maret 2016
Tanggal Persetujuan Naskah : 28 April 2016

1. PENDAHULUAN

Keamanan informasi merupakan salah satu bagian penting dalam suatu sistem informasi, karena informasi merupakan aset yang sangat berharga bagi setiap institusi ataupun perusahaan. Sedemikian pentingnya informasi sehingga informasi hanya diperuntukkan bagi orang-orang yang berkepentingan saja. Oleh karena itu, keamanan sistem informasi harus terjamin sehingga orang yang tidak berkepentingan tidak dapat mengakses informasi tersebut. Keamanan informasi tidak cukup hanya dengan pemasangan perangkat *Firewall*, tetapi dapat dioptimalkan melalui integrasi berbagai perangkat pengaman sistem informasi sehingga meminimalkan terjadinya akses oleh pengguna yang tidak berwenang.

Tindakan pengamanan informasi dari pihak-pihak yang tidak berkepentingan saat ini sudah berkembang dengan pesat, salah satunya menggunakan *Intrusion Detection System* (IDS), dimana dengan peralatan tersebut setiap ada percobaan *intrusion* oleh pihak-pihak yang tidak berkepentingan, akan ada peringatan yang dikirimkan ke pihak *administrator*. Tindakan deteksi intrusi adalah upaya untuk memantau dan mencegah berbagai upaya menyusup ke sistem dan sumber daya informasi perusahaan. Tujuan dari deteksi intrusi adalah untuk menemukan gangguan ke sumber daya informasi, dengan melakukan pengamatan berbagai kegiatan di jaringan perusahaan.

Penelitian mengenai sistem pendeteksi intrusi/*Intrusion Detection System* telah dimulai sejak tahun 1980 hingga saat ini untuk mengetahui metode pendeteksian intrusi yang performansinya lebih baik. Metode tradisional yang banyak diimplementasikan untuk mendeteksi intrusi adalah *signature-based technique*. Metode ini hanya dapat mendeteksi intrusi yang memiliki *signature* yang sesuai, sehingga *signature database* harus direvisi secara manual untuk setiap jenis intrusi yang ditemukan. Karena adanya keterbatasan ini, maka banyak penelitian yang dilakukan untuk mendeteksi intrusi dengan menggunakan teknik *data mining*. Dari hasil penelitian tersebut, banyak yang performansinya mendekati atau lebih baik jika dibandingkan dengan sistem yang tidak menggunakan teknik *data mining*.

Metode yang digunakan dalam pendeteksian intrusi menggunakan *data mining* dapat digolongkan menjadi dua bagian, yaitu *misuse detection* dan *anomaly detection*. Kelebihan dari *misuse detection* adalah mampu mendeteksi intrusi yang sudah diketahui secara akurat, tetapi tidak dapat mendeteksi jenis intrusi yang baru atau belum diketahui sedangkan *anomaly detection* dapat melakukan deteksi intrusi untuk jenis baru sebagai deviasi dari lalu lintas data yang normal.

Namun, untuk menerapkan IDS yang menggunakan *data mining* terdapat tiga kendala utama, yaitu kecenderungan hasil penelitian untuk menghasilkan *false positive* yang lebih tinggi khususnya pada saat menggunakan metode *anomaly detection*, memerlukan biaya komputasi yang tinggi (memerlukan dua tahap, yaitu pelatihan dan pengujian), serta membutuhkan data pelatihan yang jumlahnya besar.

Teknik dalam *data mining*, baik *supervised learning* maupun *unsupervised learning* dapat digunakan untuk mendeteksi intrusi pada jaringan komputer. Penelitian yang dilakukan oleh Mukkamala (2003) menggunakan metode SVM dan *Neural Network* [1]. Hasilnya, SVM memiliki akurasi yang lebih tinggi dengan waktu pengujian dan pelatihan yang lebih singkat. *Unsupervised* SVM (*One Class* SVM) memiliki akurasi yang tinggi dalam pendeteksian intrusi, tetapi *false positive* juga sangat tinggi [2]. Jadi, dari hasil penelitian tersebut, SVM merupakan teknik yang memiliki akurasi yang tinggi dalam mendeteksi intrusi dan dapat diterapkan dalam bentuk *supervised learning* maupun *unsupervised learning*.

2. LATAR BELAKANG

Walaupun SVM dapat melakukan pendeteksian intrusi dengan akurasi yang cukup tinggi, namun masih terdapat beberapa kendala dalam penerapannya. Dalam penerapan SVM belum diketahui secara pasti bagaimana pengaruh jumlah dan distribusi data yang dimasukkan ke dalam data pelatihan yang diproses terhadap performansi SVM dalam mendeteksi instrusi . Tingkat *false positive* yang cukup tinggi sangat mengganggu, karena pada saat tidak ada intrusi, IDS mengirimkan banyak peringatan ke *administrator*, demikian pula sebaliknya, pada saat akurasi pendeteksian intrusi rendah, dimungkinkan terjadi intrusi yang tidak diketahui oleh *administrator* [3].

Oleh karena itu, pada penelitian ini dilakukan penerapan IDS menggunakan *data mining*, khususnya dengan menggunakan SVM, sehingga diperoleh akurasi prediksi IDS yang terbaik, dan akurasi prediksi intrusi IDS *data mining* dengan metode SVM [4].

3. KONSEP DASAR

Berikut adalah beberapa konsep dasar terkait dengan penerapan metode SVM pada IDS dengan *Data Mining*.

a. IDS

Intrusion Detection System (IDS) adalah komponen penting pada sistem pertahanan untuk melindungi komputer dan jaringan dari aksi penyalahgunaan. Tujuan dari IDS adalah mengkarakteristikkan gejala atau kejadian yang menunjukkan terjadinya intrusi, sehingga dapat mendeteksi semua intrusi yang ada tanpa adanya kesalahan. Tujuan dari penggunaan IDS dapat bervariasi, seperti mengumpulkan informasi forensik sehingga mampu mengetahui penyusup, men-*trigger* aksi tertentu untuk melindungi sistem ketika terjadi serangan, atau digunakan sebagai alat untuk mengidentifikasi dan memperbaiki kelemahan yang terdapat pada sistem [6].

b. *Data Mining*

Data mining adalah aplikasi algoritma spesifik untuk mengekstrak pola dari data. *Data Mining* didefinisikan sebagai proses penemuan pola dalam data. *Data mining* sering juga disebut analisis data eksploratif. Data dalam jumlah besar yang diperoleh dari mesin kasir, pemindaian *barcode*, dan dari berbagai basis data dalam perusahaan, selanjutnya ditelaah, dianalisis, dihapus, dan dipakai ulang. Pencarian dilakukan pada model yang berbeda untuk memprediksi penjualan, respon pasar, keuntungan, dan lain- lain [7].

c. SVM

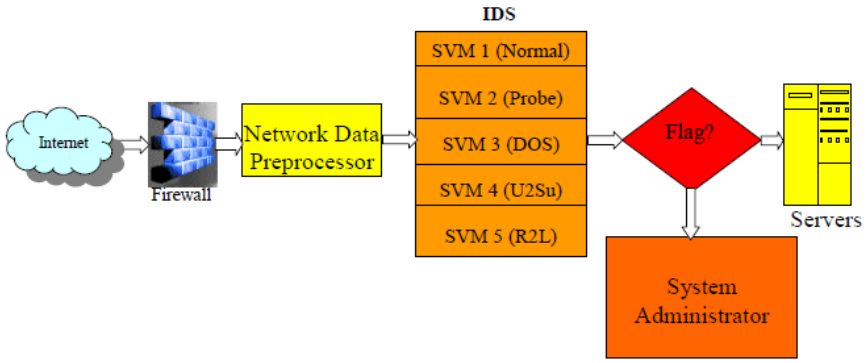
Support Vector Machine (SVM) adalah sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur (*feature space*) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan *inductive bias* yang berasal dari teori pembelajaran statistik [8]. Teori yang mendasari SVM sudah berkembang sejak tahun

1960-an, namun baru diperkenalkan oleh Vapnik, Boser, dan Guyon pada tahun 1992 dan sejak itu SVM berkembang dengan pesat. SVM adalah salah satu teknik yang relatif baru dibandingkan dengan teknik lain, tetapi memiliki performansi yang lebih baik di berbagai bidang aplikasi, seperti *bioinformatics*, pengenalan tulisan tangan, klasifikasi teks, dan lain sebagainya [9]. Proses pembelajaran pada SVM bertujuan untuk mendapatkan hipotesis berupa bidang pemisah terbaik yang tidak hanya meminimalkan *empirical risk*, yaitu rata-rata *error* pada data pelatihan, tetapi juga memiliki generalisasi yang baik [10]. Generalisasi adalah kemampuan sebuah hipotesis untuk mengklasifikasikan data yang tidak terdapat dalam data pelatihan dengan benar. Untuk menjamin generalisasi ini, SVM bekerja berdasarkan prinsip *Structural Risk Minimization* (SRM).

d. Tinjauan Penelitian Sebelumnya

1. Srinivas Mukkamala, Andrew H. Sung

   SVM mengungguli ANN dalam hal dari skalabilitas (SVM dapat melatih dengan jumlah yang lebih besar, sementara ANN akan memakan waktu lama untuk melatih atau gagal sama sekali ketika jumlah pola menjadi semakin besar) waktu pelatihan dan waktu berjalan (SVMs menjalankan pengurutan lebih cepat), serta prediksi yang lebih akurat [1]. SVM mudah mencapai akurasi deteksi yang tinggi (lebih tinggi dari 99%) untuk masing-masing dari lima kelas data, terlepas apakah semua 41 fitur digunakan, hanya fitur penting untuk setiap kelas yang digunakan, atau gabungan dari semua fitur penting untuk semua kelas digunakan. Skema jaringan yang dipergunakan pada penelitian ini dapat dilihat di gambar 1.

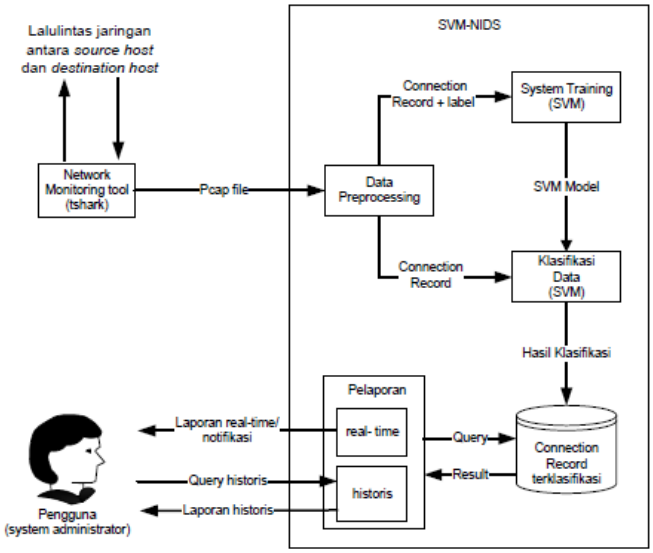


Gambar 1. Skema jaringan penelitian srinivas mukkamala, Andrew H. Sung.

2. Agustinus Jacobus, Edi Winarko

   Penerapan metode klasifikasi *support vector machine* dalam sistem deteksi intrusi yang telah dibangun dapat membantu analis dalam pembentukan profil, skenario intrusi, atau model secara otomatis, dimana dari hasil pengujian model yang dihasilkan oleh sistem ini dapat mendeteksi aksi intrusi yang dilakukan dengan tingkat akurasi dan tingkat deteksi yang tinggi, serta tingkat *false positive* yang rendah. Pengujian secara *realtime* membuktikan bahwa penerapan fungsi *complete/incomplete connection record* dalam proses *preprocessing* pembentukan *data audit* atau *connection record* dapat menjadi solusi permasalahan keterlambatan pendeteksian akibat durasi koneksi yang terlalu lama. Dari hasil pengujian secara *realtime*, sistem belum dapat mendeteksi secara efektif aksi intrusi ketika aksi tersebut baru mulai dilakukan, *connection record* yang terbentuk pada saat aksi intrusi baru dilakukan masih diklasifikasikan dengan kelas normal. Kondisi ini tidak hanya mempengaruhi tingkat pendeteksian serangan tetapi juga

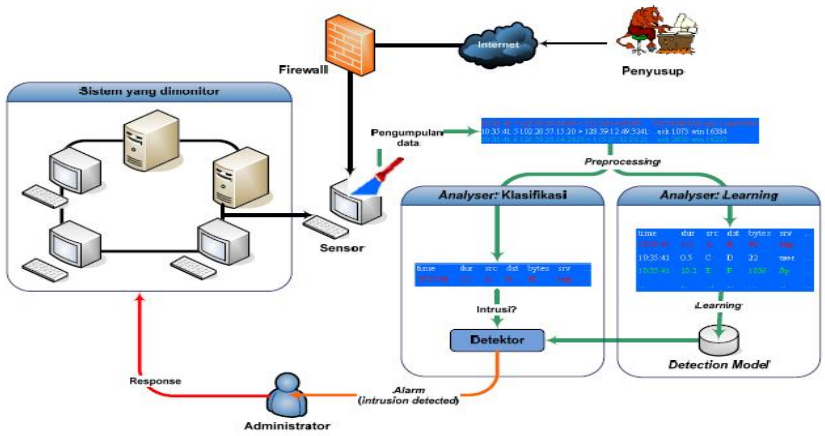
memberi celah bagi aksi-aksi intrusi yang dilakukan dengan mengirim paket data secara lambat. Untuk skema proses kerja pada penelitian ini dapat dilihat di gambar 2.



Gambar 2. Skema alur kerja penelitian Agustinus Jacobus, Edi Winarko

3. Krisantus Sembiring

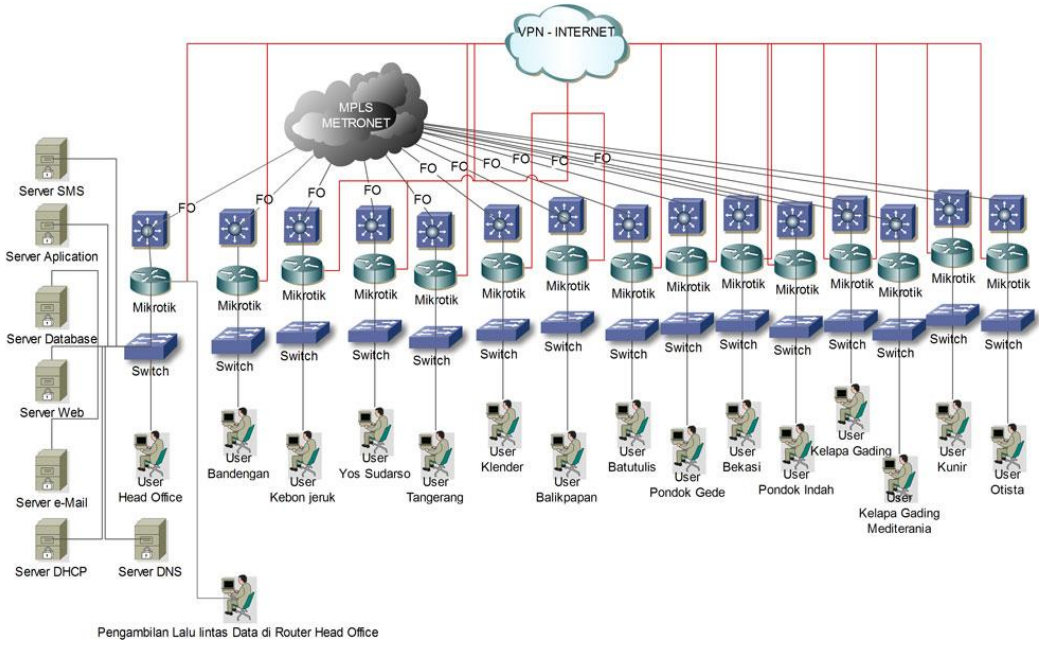
Berdasarkan hasil eksplorasi yang dilakukan pada data KDDCUP 99, model terbaik dalam mengimplementasikan SVM untuk pendeteksian intrusi pada jaringan dengan metode *misuse detection* adalah SVM *Biner*, *Multi Class* SVM *One-Against-One*, sedangkan untuk *anomaly detection* adalah *One Class* SVM dengan mengunakan data pelatihan yang seluruhnya merupakan data normal. Performansi SVM dengan *misuse detection* tidak jauh berubah pada berbagai variasi *dataset*, akan tetapi pada *anomaly detection* terjadi perubahan yang signifikan karena bergantung pada rasio data intrusi. Pada data KDDCUP 99, data untuk *One Class* SVM lebih baik dinormalisasi ke nilai maksimum dan minimumnya dengan rentang nilai [0,1], sedangkan data untuk SVM *Biner*, *Multi Class* SVM *One-Against-One* lebih baik dinormalisasi dengan data *dependent normalization* (data hasil *one-of-c encoding* juga ikut dinormalisasi). Skema jaringan dalam penelitian ini dapat dilihat di gambar 3.



Gambar 3. Skema jaringan penelitian Krisantus Sembiring.

e. Tinjauan Jaringan

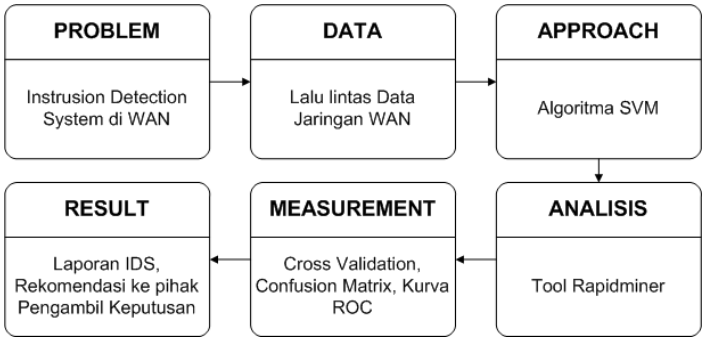
ASTRIDO Group merupakan dealer otomotif resmi terkemuka dalam bidang penjualan, perawatan, perbaikan, dan penyediaan suku cadang Toyota, Daihatsu, dan Isuzu. Jaringan WAN pada ASTRIDO Group terlihat pada gambar 4.



Gambar 4. Jaringan WAN

f. Kerangka Pemikiran

Penelitian ini terdiri atas beberapa tahap seperti terlihat pada kerangka pemikiran Gambar 5. Permasalahan (*problem*) pada penelitian ini adalah belum adanya algoritma yang akurat untuk *intrusion detection system*. Untuk itu, dibuat *approach* (model) algoritma SVM, untuk memecahkan permasalahan, untuk selanjutnya dilakukan pengujian terhadap kinerja dari metode tersebut. Pengujian menggunakan metode *Cross Validation, Confusion Matrix,* dan kurva ROC. Untuk mengembangkan aplikasi (*development*) berdasarkan model yang dibuat, digunakan *Rapid Miner*.



Gambar 5. Kerangka pemikiran

4. METODE PENELITIAN

Penelitian ini menggunakan metode penelitian eksperimen. Penelitian eksperimen melibatkan penyelidikan perlakuan pada atribut parameter atau variabel tergantung dari

peneliti dan menggunakan tes yang dikendalikan oleh si peneliti itu sendiri. Metode penelitian yang digunakan adalah sebagai berikut:

a. Pengumpulan data (*Data Gathering*)
   Pada tahap ini ditentukan data yang akan diproses, mencari data yang tersedia, memperoleh data tambahan yang dibutuhkan, mengintegrasikan semua data ke dalam *dataset*, termasuk variabel yang diperlukan dalam proses.
b. Pengolahan awal data (*Data Pre-processing*)
   Di tahap ini dilakukan penyeleksian data, data dibersihkan dan ditransformasikan ke bentuk yang diinginkan sehingga dapat dilakukan persiapan dalam pembuatan model. Tahap pengolahan awal data dilakukan untuk mempersiapkan data yang benar-benar *valid* sebelum diproses pada tahap berikutnya. Pada tahap ini dilakukan *cleansing*, transformasi, reduksi, dan seleksi fitur. Data yang didapat diolah untuk mendapatkan atribut yang relevan dan sesuai.
c. Pendekatan Analisis
   Pada tahap ini data dianalisis, dikelompokkan variabel mana yang berhubungan satu dengan lainnya. Setelah data dianalisis selanjutnya diterapkan model-model yang sesuai dengan jenis data. Pembagian data ke dalam data latihan (*training data*) dan data uji (*testing data*) juga diperlukan untuk pembuatan model.
d. Eksperimen dan pengujian model (*Model Testing and Experiment*)
   Dalam melakukan penelitian ini diperlukan eksperimen dan proses pengujian model yang diusulkan. Proses eksperimen dan pengujian model menggunakan bagian dari *dataset* yang ada. Semua *dataset* kemudian diuji dengan metode yang diusulkan pada *tools Rapid Miner*. Pengujian model berdasarkan perhitungan metode X-*Validation*.
e. Evaluasi dan validasi hasil (*Result Evaluation*)
   Pada tahap ini dilakukan evaluasi dan validasi terhadap hasil evaluasi dari eksperimen yang telah dilakukan. Model yang terbentuk akan diuji dengan menggunakan *Confusion Matrix* untuk mengetahui tingkat akurasi. *Confusion Matrix* akan menggambarkan hasil akurasi mulai dari prediksi positif yang benar, prediksi positif yang salah, prediksi *negative* yang benar, dan prediksi *negative* yang salah. Akurasi dihitung dari seluruh prediksi yang benar (baik prediksi positif maupun negatif). Semakin tinggi nilai akurasi, semakin baik pula model yang dihasilkan.

5. HASIL DAN PEMBAHASAN

Hasil ditunjukkan pada Tabel 1 untuk 10 menit pertama, 10 menit kedua, dan 10 menit ketiga.

Tabel 1. Observasi C dan ε.

Waktu	**C**	**ε**	***Accuracy***	**AUC**
10 Menit 1	0.0	0.0	86,37%	0,857
10 Menit 2	0.0	0.0	76,70%	0,858
10 Menit 3	0.0	0.0	82,94%	0,818

Hasil observasi menunjukkan bahwa nilai akurasi, yaitu 86,37 % dan AUC, yaitu 0,857 diperoleh di 10 menit pertama, 76,70 % dan AUC 0,858 diperoleh di 10 menit kedua, dan 82,94 % dan AUC 0,818 diperoleh di 10 menit ketiga. Atribut yang digunakan adalah *sourceport*, *destinationport*, dan *length* dengan *weight* dapat dilihat pada tabel 2.

Tabel 2. *Weight* atribut

Atribut	***Weight* 10 Menit 1**	***Weight* 10 Menit 2**	***Weight* 10 Menit 3**
SourcePort	1,366	0,792	1,254
DestinationPort	1,337	-1,305	1,219
Length	-0,104	0,36	-0,037

Hasil pengujian model adalah untuk mengukur tingkat akurasi dan *Area Under Curve* (AUC) dari penentuan intrusi dengan metode *cross validation*.

1. *Confusion Matrix*
   Tabel 3, Tabel 5, dan Tabel 7 menunjukkan hasil dari *confusion matrix* metode *support vector machine*.

Tabel 3. Hasil *confusion matrix* untuk 10 menit pertama

Accuracy : 86,37%		
	True Normal	*True Anomaly*
Pred. Normal	1823	351
Pred. Anomaly	0	402
Class Recall	100 %	53,39%

Berdasarkan hasil pada Tabel 3, dapat dilihat bahwa tingkat akurasi dengan menggunakan algoritma SVM adalah sebesar 86,37%, dan dapat dihitung nilai *accuracy, sensitivity, specificity,* ppv, dan npv. Hasil perhitungan terlihat pada tabel 4.

Tabel 4. Nilai *accuracy, sensitivity, specificity*, PPV, dan NPV *Metode Support Vector Machine* 10 menit pertama.

	Nilai (%)
Accuracy	86,30%
Sensitivity	83,80%
Specificity	100%
PPV	100%
NPV	53,30%

Tabel 5. Hasil *Confusion Matrix* untuk 10 menit kedua

Accuracy: 76.70%		
	True Normal	*True Anomaly*
Pred. Normal	1372	398
Pred. Anomaly	152	439
Class Recall	90,03 %	52,45%

Berdasarkan hasil pada Tabel 5, dapat dilihat bahwa tingkat akurasi dengan menggunakan algoritma SVM adalah sebesar 76,70%, dan dapat dihitung nilai *accuracy, sensitivity, specificity*, ppv, dan npv. Hasil perhitungan terlihat pada tabel 6.

Tabel 6. Nilai *accuracy, sensitivity, specificity*, PPV, dan NPV Metode *Support Vector Machine* 10 menit kedua

	Nilai (%)
Accuracy	76,70%
Sensitivity	77,51%
Specificity	74,28%
PPV	90,02%
NPV	52,44%

Tabel 7. Hasil *confusion matrix* untuk 10 menit ketiga

Accuracy: 82,94%		
	True Normal	*True Anomaly*
Pred. Normal	1718	436
Pred. Anomaly	0	402
Class Recall	100 %	47,97%

Berdasarkan hasil pada Tabel 7, dapat dilihat bahwa tingkat akurasi dengan menggunakan algoritma SVM adalah sebesar 82,94%, dan dapat dihitung nilai *accuracy*, *sensitivity, specificity*, ppv, dan npv. Hasil perhitungan terlihat pada tabel 8.

Tabel 8. Nilai *accuracy, sensitivity, specificity*, PPV, dan NPV Metode *Support Vector Machine* 10 menit ketiga

	Nilai (%)
Accuracy	82,94%
Sensitivity	79,75%
Specificity	100%
PPV	100%
NPV	47,97%

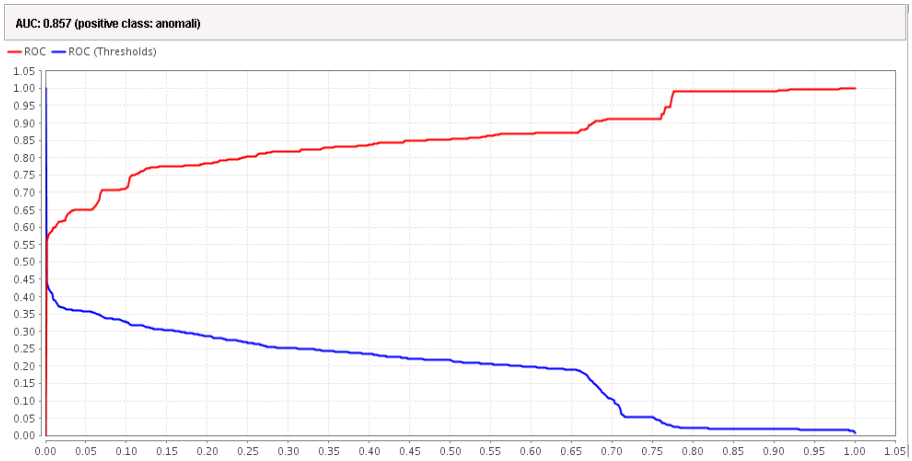
Bila digabungkan antara 10 menit pertama, kedua, dan ketiga dapat dilihat seperti pada Tabel 9.

Tabel 9. Perbandingan Nilai *accuracy, sensitivity, specificity*, PPV, dan NPV Metode *Support Vector Machine*

	10 Menit Pertama	10 Menit Kedua	10 Menit Ketiga
Accuracy	86,30%	76,70%	82,94%
Sensitivity	83,80%	77,51%	79,75%
Specificity	100%	74,28%	100%
PPV	100%	90,02%	100%
NPV	53,30%	52,44%	47,97%

2. Kurva ROC

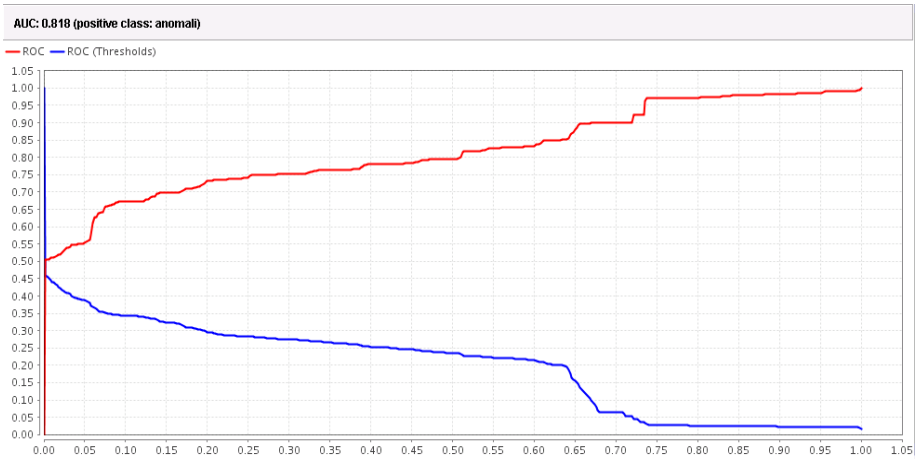
Hasil perhitungan 10 menit pertama divisualisasikan dengan kurva ROC. Gambar 6 merupakan kurva ROC untuk algoritma *Support Vector Machines* untuk 10 menit pertama. Kurva ROC pada gambar 6 mengekspresikan *confusion matrix* dari Tabel 3. Garis horizontal adalah *false positives* dan garis vertikal *true positives*.



Gambar 6. Kurva ROC 10 menit pertama

Dari Gambar 6 terlihat grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0,857.

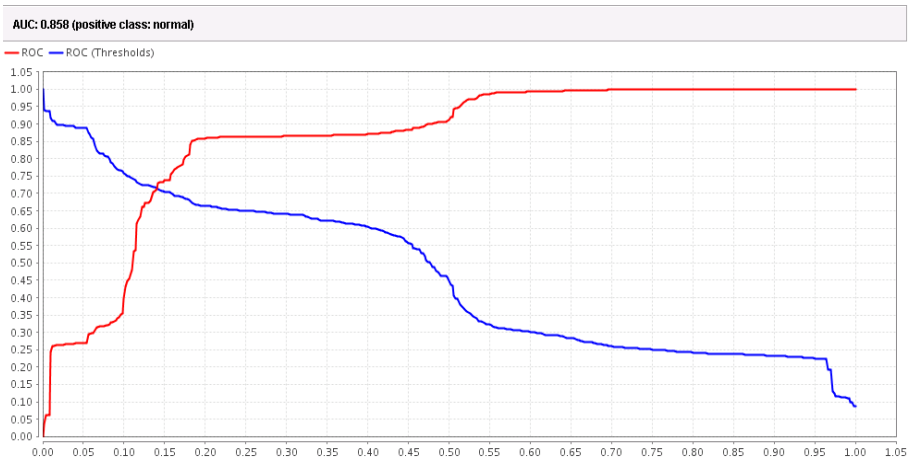
Hasil perhitungan 10 menit kedua divisualisasikan dengan kurva ROC. Gambar 7 merupakan kurva ROC untuk algoritma *Support Vector Machines* untuk 10 menit kedua. Kurva ROC pada gambar 7 mengekspresikan *confusion matrix* dari Tabel 5. Garis horizontal adalah *false positives* dan garis vertikal *true positives*.



Gambar 7. Kurva ROC 10 menit kedua

Dari Gambar 7 terlihat grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0,858.

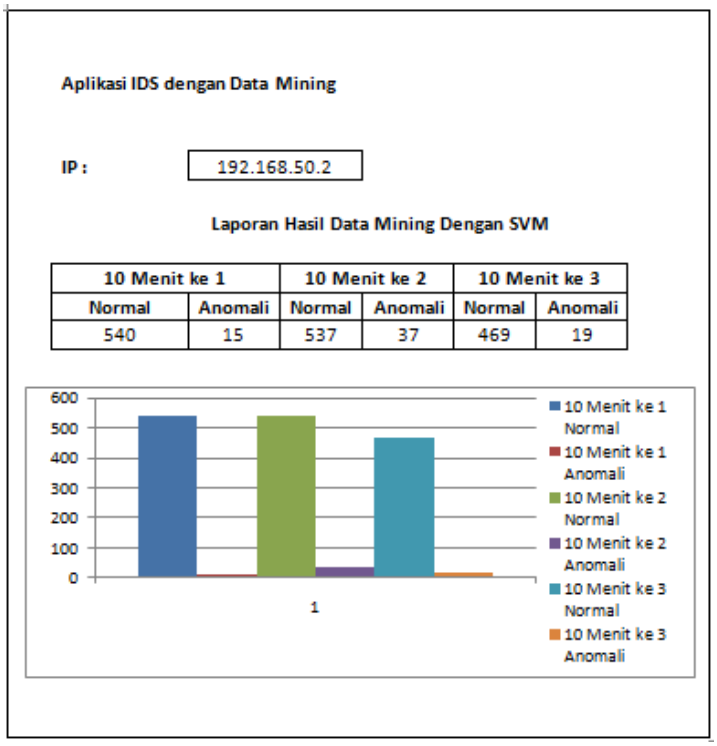
Hasil perhitungan 10 menit ketiga divisualisasikan dengan kurva ROC. Gambar 8 merupakan kurva ROC untuk algoritma *Support Vector Machines* untuk 10 menit ketiga. Kurva ROC pada gambar 8 mengekspresikan *confusion matrix* dari Tabel 7. Garis horizontal adalah *false positives* dan garis vertikal *true positives*.



Gambar 8. Kurva ROC 10 menit ketiga

Dari Gambar 8 terlihat grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0,818.

Dari hasil pengolahan data, dapat diambil prediksi untuk setiap intrusi, apakah intrusi tersebut normal atau *anomaly*, dimana apabila hasil prediksi diindikasikan sebagai *anomaly*, maka IP komputer tersebut perlu untuk diambil tindakan untuk mengurangi efek dari intrusi. Gambar 9 adalah prototipe laporan terhadap IP yang sedang di-*audit*.



Gambar 9. Contoh prototipe aplikasi IDS dengan *Data Mining*

6. KESIMPULAN

Pada penelitian ini dilakukan pengujian model dengan menggunakan *Support Vector Machines* dengan menggunakan data lalu lintas jaringan yang terkena intrusi maupun yang tidak terkena instrusi. Model yang dihasilkan diuji untuk mendapatkan nilai

accuracy, precision, recall, dan AUC dari setiap algoritma. Setelah dilakukan pengujian diperoleh hasil bahwa *support vector machine* menghasilkan nilai *accuracy* sebesar 86,30 %, nilai *Sensitivity* 83,80%, dan nilai AUC sebesar 0,857 untuk 10 menit pertama; sebesar 76,70 %, nilai *Sensitivity* 77,51%, dan nilai AUC sebesar 0,858 untuk 10 menit kedua; sebesar 82,94 %, nilai *Sensitivity* 79,75%, dan nilai AUC sebesar 0,818 untuk 10 menit ketiga. Dari hasil yang diperoleh dapat disimpulan bahwa pengujian *support vector machines* dapat memberikan pemecahan untuk permasalahan prediksi intrusi di dalam jaringan WAN secara akurat.

REFERENSI

[1]. Sung, Andrew H., Mukkamala, Srinivas. (2003). "*Feature Selection for Intrusion Detection using Neural Networks and Support Vector Machines*", TRB Annual Meeting.

[2]. Lazarevic, Aleksandar., Ertoz, Levent., Kumar, Vipin., Ozgur, Ayzel., Srivastava, Jaideep. (2003). "*A Comparative Study of Anomaly Detection Schemes in Network Intrusion Detection*".

[3]. Jacobus, Agustinus., Winarko, Edi. (2014). "Penerapan Metode Support Vector Machine pada Sistem Deteksi Intrusi secara Real-time", *IJCCS*, Vol.8, ISSN: 1978-1520.

[4]. Alpaydin, Ethem. (2010). Introduction to Machine Learning. London: The MIT Press.

[5]. Vercellis, Carlo. (2009). Business Intelligence: Data Mining and Optimization for Decision Making. New York: Springer: A John Wiley and Sons, Ltd., Publication.

[6]. Gorunescu, Florin. (2011). Data Mining: Concepts, Models, and Techniques. Verlag Berlin Heidelberg: Springer.

[7]. Sembiring, Krisantus. (2007). Penerapan Teknik Support Vectore Machine untuk Pendeteksian Intrusi pada Jaringan [Thesis]. Bandung: ITB.

[8]. Kendall, Kristoper. (1999). *A Database of Computer Attacks for the Evaluation of Intrusion Detection System* [Thesis]. MIT Lincoln Laboratory.

[9]. Kusrini, Luthfi E.T. (2009). Algoritma Data Mining. Yogyakarta: Andi Publishing.

[10]. Larose, D. T. (2005). Discovering Knowledge in Data. New Jersey: John Willey & Sons, Inc.