



TUGAS AKHIR - KI141502

EVALUASI SISTEM PENDETEKSI INTRUSI BERBASIS ANOMALI DENGAN N-GRAM DAN INCREMENTAL LEARNING

**I MADE AGUS ADI WIRAWAN
NRP 5112 100 036**

**Dosen Pembimbing I
Royyana Muslim Ijtihadie, S.Kom., M.Kom., Ph.D.**

**Dosen Pembimbing II
Baskoro Adi Pratomo, S.Kom., M.Kom.**

**JURUSAN TEKNIK INFORMATIKA
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember
Surabaya 2016**



UNDERGRADUATE THESES - KI141502

EVALUATION OF ANOMALY BASED INTRUSION DETECTION SYSTEM WITH N-GRAM AND INCREMENTAL LEARNING

**I MADE AGUS ADI WIRAWAN
NRP 5112 100 036**

**Supervisor I
Royyana Muslim Ijtihadie, S.Kom., M.Kom., Ph.D.**

**Supervisor II
Baskoro Adi Pratomo, S.Kom., M.Kom.**

**DEPARTMENT OF INFORMATICS
FACULTY OF INFORMATION TECHNOLOGY
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2016**

LEMBAR PENGESAHAN

EVALUASI SISTEM PENDETEKSI INTRUSI BERBASIS ANOMALI DENGAN N-GRAM DAN INCREMENTAL LEARNING

TUGAS AKHIR

Diajukan Guna Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada
Bidang Studi Komputasi Berbasis Jaringan
Program Studi S-1 Jurusan Teknik Informatika
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh :

I MADE AGUS ADI WIRAWAN

NRP : 5112 100 036

Disetujui oleh Dosen Pembimbing Tugas Akhir

ROYYANA MUSLIM ITHAATI

S.Kom., M.Kom., Ph.D.

NIP: 19770824 202624 1 001

BASKORO ADI PRATOMO

S.Kom., M.Kom.

NIP: 19870218 201404 1 001



**SURABAYA
JULI, 2016**

EVALUASI SISTEM PENDETEKSI INTRUSI BERBASIS ANOMALI DENGAN N-GRAM DAN INCREMENTAL LEARNING

Nama Mahasiswa : I MADE AGUS ADI WIRAWAN
NRP : 5112100036
Jurusan : Teknik Informatika FTIF-ITS
Dosen Pembimbing 1 : Royyana Muslim Ijtihadie, S.Kom.,
M.Kom., Ph.D.
Dosen Pembimbing 2 : Baskoro Adi Pratomo, S.Kom., M.Kom.

Abstrak

Keberadaan teknologi informasi yang terus berkembang dengan pesat menjadikan kebutuhan akan penggunaannya semakin hari semakin meningkat. Transaksi data melalui internet telah menjadi kebutuhan wajib hampir dari semua perangkat lunak yang ada saat ini. Perangkat lunak seperti media social, colud server, online game, aplikasi layanan pemerintah, aplikasi pengontrol suatu tempat secara remote, dsb. Tentu dengan berbagai macam penggunaan internet tersebut dibutuhkan metode untuk mengamankan jaringannya.

Sistempendeteksi intrusi atau yang pada umumnya disebut IDS (Intrusion Detection System) merupakan solusi untuk mengamankan suatu jaringan. Sistem ini nantinya bertugas untuk menentukan apakah suatu paket merupakan bentuk serangan atau paket biasa sesuai dengan kondisi tertentu. Saat ini telah banyak dikembangkan aplikasi IDS (Intrusion Detection System), namun sebagian besar yang dikembangkan berbasis signature atau menggunakan rule, dan sebagian kecil menggunakan anomali. Anomali adalah suatu metode untuk mencari penyimpangan dalam sebuah data.

Pada aplikasi ini konsep IDS yang diterapkan adalah IDS berbasis anomali dimana analisis datanya pada informasi paket data yang dikirimkan. Pada tugas akhir ini menggunakan dua metode, yaitu metode n-gram yang digunakan untuk mengitung distribusi byte karakter pada paket data sedangkan metode

mahalanonis distance digunakan untuk menghitung jarak antara paket data normal dan paket data yang berupa intrusi.

Metode mahalanobis distance dapat membedakan paket data yang normal dan paket data yang berupa intrusi dengan menghitung rata-rata dan standar deviasi dari paket data.

Kata kunci : N-Gram, Mahalanobis Distance, Incremental Learning

EVALUATION OF ANOMALY BASED INTRUSION DETECTION SYSTEM WITH N-GRAM AND INCREMENTAL LEARNING

Student's Name : I MADE AGUS ADI WIRAWAN
Student's ID : 5112100036
Department : Teknik Informatika FTIF-ITS
First Advisor : Royyana Muslim Ijtihadie, S.Kom.,
M.Kom., Ph.D.
Second Advisor : Baskoro Adi Pratomo, S.Kom.,
M.Kom.

Abstract

The rapid development of information technology is inevitable which made its necessity is growing every single day. Data transaction through internet has become the primary need of most software nowadays. Software like social media, cloud server, online game, e-government, remote application, etc. With the various needs of the internet, it is obvious that we need a method that can guarantee its safety.

IDS which stands for Intrusion Detection System is the solution to protect the internet network. This system will decide whether a packet is safe or dangerous for the network depends on certain condition. Nowadays many IDS (Intrusion Detection System) has been developed, but most are developed base signature or use the rule, and a small part sing anomaly. Anomaly is a method to look for irregularities in the data.

In this application IDS concept that is applied is based anomaly in which the data analysis on the data packets transmitted. In this thesis using two methods, the n-gram method used to calculate the distribution of byte character data paket while the mahalanobis distance methods used to calculated the distance between the normal data packets and intrusion data packets.

Mahalanobis distance methods can distinguish between normal data packets and intrusion data packets by calculating the average and standar deviation of the data packets.

Keyword : N-Gram, Mahalanobis Distance, Incremental Learning

DAFTAR ISI

LEMBAR PENGESAHAN.....	v
Abstrak.....	viii
Abstract.....	x
KATA PENGANTAR.....	xii
DAFTAR ISI.....	xiv
DAFTAR GAMBAR.....	xviii
DAFTAR TABEL.....	xx
DAFTAR PERSAMAAN.....	xxii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	3
1.4 Tujuan.....	3
1.5 Manfaat.....	3
1.6 Metodologi.....	3
1.7 Sistematika Penulisan Laporan Tugas Akhir.....	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 IDS.....	7
2.2 IDS Berbasis Anomali.....	8
2.3 Jpcap.....	8
2.4 N-Gram.....	10
2.5 Simplified Mahalanobis Distance.....	11
2.6 <i>Incremental Learning</i>	13
2.7 DARPA 1999.....	14
2.7.1 Arsitektur Simulasi DARPA 1999.....	15
2.7.2 Jenis – jenis Serangan dari DARPA 1999.....	16
BAB III DESAIN DAN PERANCANGAN.....	19
3.1 Deskripsi Umum Sistem.....	19
3.2 Perancangan.....	20
3.2.1 Alur Kerja Sistem Secara Umum.....	20
3.2.2 Perancangan Arsitektur Jaringan.....	22
3.2.3 Perancangan Proses <i>Training</i> Data Set.....	22
3.2.4 Perancangan Proses <i>Sniffing</i>	23

3.2.5	Perancangan Proses Identifikasi Intrusi.....	25
3.2.6	Rancangan Antarmuka	26
3.2.7	Rancangan Luaran Sistem.....	26
BAB IV IMPLEMENTASI		27
4.1	Lingkungan Implementasi	27
4.1.1	Perangkat Lunak	27
4.1.2	Perangkat Keras.....	27
4.2	Implementasi Proses	28
4.2.1	Data set.....	28
4.2.2	Implementasi Proses Rekonstruksi Paket Data	29
4.2.3	Implementasi Proses Penggunaan Metode <i>N-Gram</i>	30
4.2.4	Implementasi Perancangan Model Data <i>Training</i>	31
4.2.5	Implementasi <i>Sniffer</i>	32
4.2.6	Implementasi Proses Penggunaan Metode Mahalanobis <i>Distance</i>	32
4.2.7	Implementasi Pendeteksian Intrusi	33
4.2.8	Implementasi Proses <i>Incremental Learning</i>	33
BAB V UJI COBA DAN EVALUASI		35
5.1	Lingkungan Uji Coba	35
5.2	Skenario Uji Coba.....	37
5.2.1	Uji Fungsionalitas	37
5.2.1.1	Uji Coba pengguna normal mengakses <i>server</i>	38
5.2.1.2	Uji Coba Proses Rekonstruksi Paket Data	39
5.2.1.3	Uji Coba Proses Menghitung N-Gram Paket Data....	41
5.2.1.4	Uji Coba Proses Membuat Model Data <i>Training</i>	43
5.2.1.5	Uji Coba Sniffing	44
5.2.1.6	Uji Coba Proses Menghitung Jarak Mahalanobis	45
5.2.1.7	Uji Coba Proses Deteksi Paket Data Normal dan Paket Data Intrusi.....	47
5.2.1.8	Uji Coba Proses <i>Incremental Learning</i>	48
5.2.2	Uji Coba Performa	50
5.2.2.1	Uji Coba Performa Sistem.....	51
5.2.2.2	Uji Coba Kecepatan Pendeteksian	59
5.2.2.3	Uji Coba Akurasi	60
BAB VI KESIMPULAN DAN SARAN.....		71

6.1 Kesimpulan	71
6.2 Saran	71
DAFTAR PUSTAKA	73
LAMPIRAN	75
A. Kode Sumber	75
A.1. Kode Sumber Proses Rekonstruksi Paket Data.....	75
A.2. Kode Sumber Proses Penggunaan Metodel N-Gram.....	81
A.3. Kode Sumber Proses Perancangan Model Data Training ...	82
A.4. Kode Sumber Sniffing	86
A.5. Kode Sumber Proses Penggunaan Metode Mahalanobis Distance.....	89
A.6. Kode Sumber Proses Pendeteksian Serangan.....	90
A.7. Kode Sumber Proses Incremental Learning	94
BIODATA PENULIS.....	97

DAFTAR GAMBAR

Gambar 2.1 Contoh penggunaan Jpcap.....	9
Gambar 2.2 Kode sumber penggunaan Jpcap untuk <i>offline capture</i>	9
Gambar 2.3 Contoh keluaran <i>offline capture</i>	10
Gambar 2.4 Arsitektur DARPA 1999.....	16
Gambar 3.1 Diagram Alir kerja sistem secara umum.....	21
Gambar 3.2 Topologi jaringan yang akan digunakan.....	22
Gambar 3.3 Proses training data set.....	23
Gambar 3.4 Proses <i>sniffing</i>	24
Gambar 3.5 Proses Identifikasi Intrusi.....	25
Gambar 3.6 Contoh <i>file</i> konfigurasi	26
Gambar 3.7 Contoh log hasil luaran sistem.....	26
Gambar 4.1 <i>Pseudocode</i> untuk Rekonstruksi paket data	30
Gambar 4.2 <i>Pseudocode</i> untuk menghitung N-Gram paket data.....	30
Gambar 4.3 <i>Pseudocode</i> untuk membuat model data <i>training</i> ...	31
Gambar 4.4 <i>Pseudocode</i> untuk <i>sniffer</i>	32
Gambar 4.5 <i>Pseudocode</i> penggunaan metode Mahalanobis Distance.....	33
Gambar 4.6 <i>Pseudocode</i> untuk Pendeteksian Intrusi.....	33
Gambar 4.7 <i>Pseudocode</i> untuk <i>Incremental Learning</i>	34
Gambar 5.1 Luaran yang dihasilkan oleh komputer pengkases normal dengan IP:192.168.57.2	38
Gambar 5.2 Luaran yang dihasilkan oleh komputer penyerang dengan IP:192.168.57.3.....	39
Gambar 5.3 Potongan hasil paket data tanpa rekonstruksi.....	40
Gambar 5.4 Potongan hasil paket data setelah direkonstruksi.....	41
Gambar 5.5 Potongan hasil N-Gram paket data.....	42
Gambar 5.6 Potongan hasil model data <i>training</i>	44
Gambar 5.7 Potongan hasil <i>sniffing</i>	45
Gambar 5.8 Potongan hasil menghitung jarak mahalanobis.....	46
Gambar 5.9 Potongan hasil deteksi paket data normal dan paket data berupa intrusi	48

Gambar 5.10 Potongan hasil data sebelum proses <i>incremental learning</i>	49
Gambar 5.11 Potongan hasil data setelah proses <i>incremental learning</i>	50
Gambar 5.12 HTOP CPU ketika sistem belum berjalan	51
Gambar 5.13 HTOP CPU ketika <i>training</i> data set berjalan.....	51
Gambar 5.14 HTOP CPU ketika identifikasi berjalan	52
Gambar 5.15 Grafik persentase utilisasi CPU.....	52
Gambar 5.16 HTOP RAM ketika sistem belum berjalan	53
Gambar 5.17 HTOP RAM ketika <i>training</i> data set berjalan.....	53
Gambar 5.18 HTOP RAM ketika identifikasi berjalan.....	53
Gambar 5.19 Grafik persentase utilisasi RAM	54
Gambar 5.20 Tampilan halaman web yang akan diakses.....	55
Gambar 5.21 Luaran ApacheBench untuk skenario 1	56
Gambar 5.22 Luaran ApacheBecnh untuk skenario 2.....	57
Gambar 5.23 Luaran ApacheBench untuk skenario 3.....	58
Gambar 5.24 Grafik waktu akses web	59
Gambar 5.25 Grafik durasi waktu pendeteksian intrusi	60
Gambar 5.26 Model Confussion Matrix untuk pengujian	62

DAFTAR TABEL

Tabel 2.1 Format data kasar didalam Mahalanobis <i>Distance</i>	12
Tabel 4.1 Data set file Paket Data	29
Tabel 4.2 Daftar bagian paket yang dibutuhkan program	30
Tabel 5.1 Prosedur pengguna normal mengakses <i>server</i>	38
Tabel 5.2 Prosedur rekonstruksi paket data	39
Tabel 5.3 Prosedur menghitung N-Gram paket data	41
Tabel 5.4 Prosedur membuat model data <i>training</i>	43
Tabel 5.5 Prosedur <i>sniffing</i>	44
Tabel 5.6 Prosedur menghitung jarak mahalanobis	46
Tabel 5.7 Prosedur deteksi paket data normal dan paket data berupa intrusi	47
Tabel 5.8 Prosedur proses <i>incremental learning</i>	49
Tabel 5.9 Metode akses komputer penyerang.....	60
Tabel 5.10 Data uji	61
Tabel 5.11 Hasil Uji Data <i>Training</i> minimum jarak paket data intrusi	64
Tabel 5.12 Hasil Uji Data <i>Training</i> maksimum jarak paket data normal	64
Tabel 5.13 Threshold untuk masing-masing port.....	65
Tabel 5.14 Hasil Uji Data <i>Testing</i> minggu ke-5 tanpa proses <i>incremental learning</i>	65
Tabel 5.15 <i>Confussion matrix</i> uji coba 1a.....	65
Tabel 5.16 Hasil penilaian percobaan 1a dengan ukuran window 10000	66
Tabel 5.17 hasil penilaian percobaan 1a dengan ukuran window 20000	66
Tabel 5.18 Hasil Uji Data <i>Testing</i> minggu ke-5 dengan proses <i>incremental laeraning</i>	66
Tabel 5.19 <i>Confussion matrix</i> uji coba 1b.....	67
Tabel 5.20 Hasil penilaian percobaan 1b dengan ukuran window 10000	67
Tabel 5.21 hasil penilaian percobaan 1b dengan ukuran window 20000	67

Tabel 5.22 Skenario serangan	68
Tabel 5.23 Hasil Uji Data <i>Testing</i> secara <i>real-time</i>	68
Tabel 5.24 <i>Confussion matrix</i> uji coba 2a	68
Tabel 5.25 Hasil penilaian percobaan 2a FTP <i>brute force</i>	69
Tabel 5.26 Hasil penilaian percobaan 2a Telnet <i>brute force</i>	69
Tabel 5.27 Hasil Uji Data <i>Testing</i> secara <i>real-time</i>	69
Tabel 5.28 <i>Confussion matrix</i> uji coba 2b	70
Tabel 5.29 Hasil penilaian percobaan 2b FTP <i>brute force</i>	70
Tabel 5.30 Hasil penilaian percobaan 2b Telnet <i>brute force</i>	70

DAFTAR PERSAMAAN

Persamaan 2.1	12
Persamaan 2.2	13
Persamaan 2.3	13
Persamaan 2.4	14
Persamaan 2.5	14
Persamaan 5.1	61
Persamaan 5.2	63
Persamaan 5.3	63
Persamaan 5.4	63
Persamaan 5.5	63
Persamaan 5.6	63
Persamaan 5.7	64

BAB VI

KESIMPULAN DAN SARAN

Pada bab ini akan dibahas mengenai kesimpulan yang dapat diambil dari perancangan sistem hingga hasil pengujian. Selain itu juga akan dibahas mengenai hasil yang sudah dicapai dan belum dicapai. Pada bab ini juga akan menjawab pertanyaan yang dikemukakan pada Bab 1. Pada penutup ini juga terdapat saran-saran untuk pengembangan selanjutnya.

6.1 Kesimpulan

Dalam proses pengerjaan Tugas Akhir yang melalui tahap perancangan, implementasi, serta uji coba, didapatkan kesimpulan sebagai berikut :

1. Metode Mahalanobis *Distance* dapat digunakan untuk mengklasifikasikan antara paket data normal dan paket data yang berupa intrusi, namun tidak berlaku untuk protokol HTTP.
2. Sistem yang dibuat untuk pendeteksi intrusi menggunakan metode Mahalanobis *Distance* tanpa proses *incremental learning* dapat mendeteksi intrusi dengan persentase kebenaran sekitar 93%, namun dengan tambahan proses *incremental learning* hanya dapat mendeteksi intrusi dengan persentase kebenaran sekitar 20%. Dari hasil tersebut, dengan tambahan proses *incremental learning* mengurangi tingkat akurasi pendeteksian intrusi.

6.2 Saran

Adapun saran-saran yang diberikan untuk pengembangan sistem ini selanjutnya adalah karena membedakan paket data normal dengan paket data serangan menggunakan metode mahalanobis *distance* dengan proses *incremental learning* kurang akurat dibandingkan tanpa proses *incremental learning*. Hal ini dikarenakan dengan menambahkan proses *incremental learning*,

rata-rata dan standar deviasi pada model diperbaharui tetapi *threshold* yang digunakan untuk mendeteksi intrusi tidak diperbaharui, sehingga *threshold* yang ada tidak akurat untuk mendeteksi intrusi. Perlu ada implementasi metode lain sehingga dapat membantu meningkatkan keakuratan pendeteksian intrusi.

DAFTAR PUSTAKA

- [1] SANS Institute, "Understanding Intrusion Detection System," *SANS Institute Reading Room*, pp. 1-9, 2001.
- [2] "Intrusion detection system," [Online]. Available: https://en.wikipedia.org/wiki/Intrusion_detection_system. [Diakses 22 June 2016].
- [3] K. Fuji, "a Java library for capturing and sending network packets," Jpcap, 15 May 2007. [Online]. Available: <http://jpcap.gitspot.com/>. [Diakses 23 May 2016].
- [4] A. Hanafi, "Pengenalan Bahasa Suku Bangsa Indonesia Berbasis Teks Menggunakan Metode N-gram. IT TELKOM," 2009.
- [5] "Mahalanobis distance," [Online]. Available: https://en.wikipedia.org/wiki/Mahalanobis_distance. [Diakses 22 June 2016].
- [6] D. E. Knuth, "The Art of Computer Programming," *Fundamental Algorithms*. Addison Wesley, vol. 1, 1973.
- [7] MIT Lincoln Laboratory, "MIT Lincoln Laboratory: Cyber system & technolog: DARPA Intrusion Detection," MIT Lincoln Laboratory, [Online]. Available: https://www.ll.mit.edu/mission/communications/cyber/CST_corpora/ideval/docs/index.html. [Diakses 23 Mei 2016].
- [8] V. Galleys, "Cross Validation," 2006. [Online]. Available: http://www.cse.iitb.ac.id/~tarung/smt/papers_ppt/ency-cross-validation.pdf. [Diakses 24 June 2016].
- [9] Kohavi, "Confusion Matrix," 1999. [Online]. Available: http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html. [Diakses 24 June 2016].

BIODATA PENULIS



I Made Agus Adi Wirawan, lahir di Desa Celagi, 4 Agustus 1994. Penulis adalah anak kedua dari dua bersaudara. Menempuh pendidikan di SD No. 3 Denbantas, SMP Negeri 1 Tabanan, SMA Negeri 1 Tabanan, dan terakhir melanjutkan kuliah di jurusan Teknik Informatika – ITS.

Selama berkuliah penulis aktif dalam kegiatan dan organisasi keprofesian informatika sebagai administrator sekaligus koordinator laboratorium Arsitektur dan

Jaringan Komputer. Pernah mengikuti dan mendapatkan sertifikasi HCNA-WCDMA yang diselenggarakan oleh Huawei.

Selain itu, penulis juga aktif dalam organisasi kampus sebagai anggota Himpunan Mahasiswa Teknik Computer-Informatika, staf departemen dalam negeri dan menjadi panitia berbagai kegiatan di tingkat jurusan maupun fakultas.

Selain menjalankan tugas mahasiswa, penulis juga aktif menjadi asisten dosen mata kuliah sistem operasi, jaringan komputer, dan keamanan informasi dan jaringan, disamping asisten dosen juga sekaligus menjadi asisten praktikum untuk mata kuliah sistem operasi dan jaringan komputer,

Ketertarikan penulis dibidang informatika berada pada bidang sistem teknologi informasi, sekuritas jaringan, perancangan keamanan sistem dan jaringan, teknologi antar jaringan, dan teknologi tepat guna.

Penulis dapat dihubungi dengan mengirimkan pesan elektronik ke alamat imadeagus.04@gmail.com.