



MEMPREDIKSI SERANGAN PADA SIM (*SECURITY INFORMATION MANAGEMENT*) DENGAN MENGGUNAKAN ALGORITMA *HIDDEN MARKOV MODEL*

SKRIPSI

**RICO ANDREAS
1610511080**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
2020**



MEMPREDIKSI SERANGAN PADA SIM (*SECURITY INFORMATION MANAGEMENT*) DENGAN MENGGUNAKAN ALGORITMA *HIDDEN MARKOV MODEL*

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
Sarjana Komputer**

**RICO ANDREAS
1610511080**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
2020**

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Rico Andreas

NIM : 1610511080

Tanggal : 18 Mei 2018

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 20 Mei 2020

Yang Menyatakan,



PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini.

Nama : Rico Andreas

NIM : 1610511080

Fakultas : Ilmu Komputer

Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Memprediksi Serangan pada SIM (*Security Information Management*) dengan

Menggunakan *Hidden Markov Model*

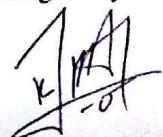
Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 20 Mei 2020

Yang menyatakan,



(Rico Andreas)

LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa Skripsi berikut:

Nama : Rico Andreas

NIM : 1610511080

Program Studi : Informatika

Judul Tugas Akhir : Memprediksi Serangan Pada SIM (*Security Information Management*) Dengan Menggunakan *Hidden Markov Model*

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Yuni Widiastiwi, S.Kom., M.Si.

Penguji I

Mayanda Mega Santoni, S.Kom., M.Kom.

Penguji II

Henki Bayu Seta, S.Kom., MTI.

Pembimbing I



Dr. Ermawita, M. Kom.

Dekan

Nurul Chamidah, S.Kom, M.Kom.

Pembimbing II

Anita Muliawati, S.Kom., MTI.

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 19 Juni 2020



MEMPREDIKSI SERANGAN PADA SIM (*SECURITY INFORMATION MANAGEMENT*) DENGAN MENGGUNAKAN ALGORITMA HIDDEN MARKOV MODEL

Rico Andreas

1610511080

Abstrak

Website merupakan suatu aplikasi yang mudah diakses dimanapun dan kapanpun, dalam kemudahan tersebut terdapat suatu serangan yang dapat dilakukan pada *website* yang tertuju langsung pada *web server*. Serangan-serangan tersebut memiliki ancaman yang sudah didata oleh OWASP (*Open Web Application Security Project*) pada tahun 2017 sehingga menciptakan informasi yang ada pada OWASP Top 10 Security – 2017 yang khusus pada aplikasi *web*. Dengan ancaman-ancaman tersebut penelitian ini dilakukan untuk membuat sistem yang dapat mendeteksi suatu serangan yang terjadi pada *website* khususnya pada *web server* dan dapat memantau serta menampilkan informasi kegiatan yang ada pada *web server* dengan *client*. *Security Information Management* (SIM) akan membaca data *access log* dan *error log* yang telah dicatat oleh *web server* lalu data tersebut akan dilakukan *training* dan *testing* dengan menggunakan algoritma *Hidden Markov Model* sehingga mendapat suatu model *learning* bagi sistem untuk mendeteksi sebuah serangan, serta *access log* dan *error log* akan diterjemahkan menjadi suatu informasi yang mudah dibaca oleh *sysadmin* kedalam suatu *dashboard*. Penelitian ini diharapkan dapat menghasilkan suatu model yang dapat mendeteksi sekaligus memantau kegiatan *web server* dalam sebuah serangan.

Kata kunci: *Access log, Error log, Security Information Management, SIM, Hidden Markov Model*

**PREDICTING AN ATTACK TO SIM (SECURITY
INFORMATION MANAGEMENT) USING HIDDEN MARKOV
ALGORITHM MODEL**

Rico Andreas

1610511080

Abstract

Website is an app that is easily accessed anywhere and anytime, in the ease there is an attack by hackers that can be carried out on a website that is directed directly to the web server. These attacks have threats that have been recorded by OWASP (Open Web Application Security Project) in 2017 thus creating information on OWASP Security 10 - 2017 specifically on web applications. With these threats, this research was conducted to create a system that can detect an attack that occurs on a website, especially on a web server and can monitor and display information on existing activities on a web server with a client. Security Information Management (SIM) will read access log and error log data that have been recorded by the web server and then the data will be conducted training and testing using the Hidden Markov Model algorithm so that it gets a learning model for the system to detect an attack, as well as access logs and error log will be translated into information that is easily read by sysadmin into a dashboard. This research is expected to produce a model that can detect and monitor web server activities in an attack.

Keywords: Access log, Error log, Security Information Management, SIM, Hidden Markov Model

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala karunia-Nya, sehingga Skripsi ini berhasil diselesaikan. Penulis ingin mengucapkan terima kasih kepada:

1. Ibu tercinta penulis yaitu Murni Elfrida Manik yang terus selalu mendoakan, memberikan dorongan dan nasihat yang terbaik agar dapat menyelesaikan skripsi ini.
2. Bapak Henki Bayu Seta, M.Kom., MTI selaku dosen pembimbing I Skripsi yang membantu memberikan saran yang bermanfaat.
3. Ibu Nurul Chamidah, S.Kom., M.Kom selaku dosen pembimbing II Skripsi yang membantu memberikan pembelajaran dan saran yang bermanfaat.
4. Hidayah Khusnul Hotimah Darnisa Azzahra Nasution, Nadya Permatasari, dan Suci Dilasari Kamil yang memberikan *support* dan membantu menyelesaikan Skripsi ini.
5. Ibu, Bapak Dosen Informatika UPN Veteran Jakarta atas segala pembelajaran dan ilmu-ilmu yang bermanfaat semasa perkuliahan.
6. Teman-teman Informatika 2016 yang selalu mendukung dalam menyelesaikan Skripsi ini.
7. Senior Informatika 2015 dan 2014 yang memberikan dukungan moril semasa perkuliahan.

Akhir kata, semoga skripsi ini dapat bermanfaat bagi para pembacanya.

Jakarta, 18 Mei 2019

Penulis,

Rico Andreas

DAFTAR ISI

MEMPREDIKSI SERANGAN PADA SIM (<i>SECURITY INFORMATION MANAGEMENT</i>) DENGAN MENGGUNAKAN ALGORITMA <i>HIDDEN MARKOV MODEL</i>	ii
PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI	iv
LEMBAR PERSETUJUAN.....	v
Abstrak	vi
Abstract	vii
KATA PENGANTAR	viii
Daftar Tabel	xii
Daftar Gambar.....	xiii
Daftar Simbol	xiv
BAB 1 Pendahuluan.....	1
1.1 Latar Belakang	1
1.1 Rumusan Masalah	3
1.2 Tujuan Penelitian.....	3
1.3 Manfaat Penelitian.....	3
1.4 Batasan Masalah.....	4
1.5 Luaran yang Diharapkan	4
1.6 Sistematika Penulisan.....	4
BAB 2 Tinjauan Pustaka.....	6
2.1 <i>Website</i>	6
2.2 <i>Web Server</i>	6
2.3 HTTP/HTTPS.....	7
2.4 <i>HTTP Status Code</i>	7
2.4.1 <i>Informational Status Code (1xx)</i>	8
2.4.2 <i>Client Request Successful (2xx)</i>	8
2.4.3 <i>Request Redirected (3xx)</i>	8

2.4.4	<i>Client Request Incomplete (4xx)</i>	8
2.4.5	<i>Server Errors (5xx)</i>	9
2.5	OWASP	9
2.5.1	<i>Injection</i>	10
2.5.2	<i>Broken Authentication</i>	10
2.5.3	<i>Sensitive Data Exposure</i>	10
2.5.4	<i>XML External Entities (XXE)</i>	11
2.5.5	<i>Broken Access Control</i>	11
2.5.6	<i>Security Misconfiguration</i>	11
2.5.7	<i>Cross-site Scripting (XSS)</i>	11
2.5.8	<i>Insecure Deserialization</i>	12
2.5.9	<i>Using Components with Known Vulnerabilities</i>	12
2.5.10	<i>Insufficient Logging & Monitoring</i>	12
2.6	<i>Log</i>	12
2.6.1	<i>Error log</i>	13
2.6.2	<i>Access log</i>	14
2.7	<i>JSON</i>	15
2.8	<i>Bash (Bourne Again Shell)</i>	16
2.9	<i>ELK (Elasticsearch Logstash Kibana)</i>	17
2.10	<i>Security Information Management (SIM)</i>	18
2.11	<i>Hidden Markov Model</i>	19
2.11.1	<i>Algoritma Viterbi</i>	22
2.12	<i>Laplacian Smoothing</i>	23
2.13	<i>Studi Literatur</i>	23
BAB 3 Metodologi Penelitian	26
3.1	Tahapan Penelitian	26
3.2	Metode Penelitian.....	27
3.2.1	<i>Identifikasi Masalah</i>	27
3.2.2	<i>Perumusan Masalah</i>	27
3.2.3	<i>Studi Literatur</i>	27
3.2.4	<i>Pengumpulan Data</i>	27
3.2.5	<i>Pra-proses data</i>	27
3.2.6	<i>Modeling System</i>	28

3.2.7	Hasil	30
3.2.8	Evaluasi	30
3.2.9	<i>Dashboard</i>	30
3.3	Alat Bantu Penelitian.....	30
3.4	Jadwal Penelitian.....	31
BAB 4		33
4.1	Pengumpulan Data	33
4.2	Pra-proses data	36
4.3	<i>Modeling System</i>	43
4.4	Hasil.....	64
4.5	Evaluasi	68
4.6	<i>Dashboard</i>	70
BAB 5		78
5.1	Kesimpulan.....	78
5.2	Saran.....	78
DAFTAR PUSTAKA		80
RIWAYAT HIDUP.....		83
LAMPIRAN.....		84

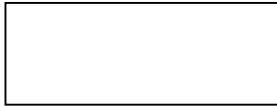
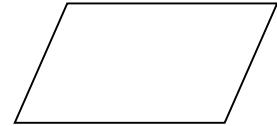
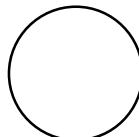
Daftar Tabel

Tabel 1.1 Klasifikasi Ancaman pada tahun 2017.....	1
Tabel 3.1 Jadwal Penelitian.....	32
Tabel 4.1 Webserverlog Setelah Integrasi Dalam Bentuk <i>Dataset</i> Format .csv....	39
Tabel 4.2: <i>Dataset</i> Webserverlog Setelah Dilakukan Klasifikasi Secara Manual.	41
Tabel 4.3: <i>Transition Probability</i> Untuk <i>Train</i> 70% dan <i>Test</i> 30%.....	46
Tabel 4.4: <i>Transition Probability</i> Untuk <i>Train</i> 80% dan <i>Test</i> 20%.....	47
Tabel 4.5: <i>Transition Probability</i> Untuk <i>Train</i> 90% dan <i>Test</i> 10%.....	47
Tabel 4.6: Sampel <i>Data Testing</i>	53
Tabel 4.7 Pencarian probabilitas terbesar skenario 1 kelompok akses pertama	56
Tabel 4.8 Pencarian probabilitas terbesar skenario 1 kelompok akses kedua.....	57
Tabel 4.9 Pencarian probabilitas terbesar skenario 2 kelompok akses pertama	59
Tabel 4.10 Pencarian probabilitas terbesar skenario 2 kelompok akses kedua.....	61
Tabel 4.11 Pencarian probabilitas terbesar skenario 3 kelompok akses pertama ..	63
Tabel 4.12 Pencarian probabilitas terbesar skenario 3 kelompok akses kedua.....	64
Tabel 4.13: Hasil prediksi skenario 1.....	65
Tabel 4.14: Hasil prediksi skenario 2.....	66
Tabel 4.15: Hasil prediksi skenario 3.....	67
Tabel 4.16: <i>Confusion Matrix</i> Untuk <i>Train</i> 70% dan <i>Test</i> 30%	68
Tabel 4.17: <i>Confusion Matrix</i> Untuk <i>Train</i> 80% dan <i>Test</i> 20%	68
Tabel 4.18: <i>Confusion Matrix</i> Untuk <i>Train</i> 90% dan <i>Test</i> 10%	69
Tabel 4.19 Nilai Hasil Akurasi, Presisi, dan <i>Recall</i>	69

Daftar Gambar

Gambar 2.1 <i>Error Log</i>	13
Gambar 2.2: <i>Access Log</i>	15
Gambar 2.3: Diagram ELK Stack	17
Gambar 2.4: <i>Markov Chain</i>	19
Gambar 2.5: Representasi <i>variable state HMM</i>	21
Gambar 3.1 <i>Flowchart</i> Tahapan Penelitian	26
Gambar 3.2: <i>Flowchart</i> pra-proses data.....	28
Gambar 3.3 Alur Perancangan Sistem	29
Gambar 4.1 <i>Flowchart</i> Integrasi <i>Access Log</i> dan <i>Error Log</i>	38
Gambar 4.2: Model HMM Yang Dibentuk Untuk Webserverlog	44
Gambar 4.3: <i>Flowchart</i> Menghitung <i>Transition Probability</i>	45
Gambar 4.4: <i>Flowchart</i> Menghitung <i>Emission Probability</i>	48
Gambar 4.5: <i>Flowchart</i> Implementasi <i>Laplacian Smoothing</i>	51
Gambar 4.6 Diagram Akurasi, Presisi, dan <i>Recall</i> HMM Semua Skenario	69
Gambar 4.7: Tampilan <i>Discover</i> pada data <i>log Apache Server</i>	71
Gambar 4.8: Titik Koordinat <i>user</i> dalam mengakses SIAKAD	72
Gambar 4.9: Titik Koordinat <i>user</i> dalam mengakses SIAKAD luar Indonesia.....	73
Gambar 4.10: Negara dengan perangkat untuk mengakses SIAKAD	74
Gambar 4.11: Negara luar Indonesia dengan perangkat mengakses SIAKAD	74
Gambar 4.12: <i>IP Address</i> dengan <i>request</i> pada SIAKAD	75
Gambar 4.13: Visualisasi data tabel untuk <i>report</i>	76
Gambar 4.14: <i>Dashboard</i> SIAKAD SIM	77

Daftar Simbol

Simbol	Nama Simbol	Keterangan
	Simbol Proses	Menggambarkan Proses
	Simbol Dokumen	Dokumen yang dibutuhkan dalam proses sistem
	Simbol arah data atau arus data	Sebagai petunjuk arah data dan arus data pada proses
	Simbol Terminator	Simbol untuk permulaan atau akhir dari suatu kegiatan
	Simbol Data	Simbol sebagai masukan atau keluaran data untuk suatu proses
	Simbol konektor	Simbol untuk sambungan pada halaman yang sama