PENDETEKSIAN *MALWARE* ANDROID MENGGUNAKAN PENGGABUNGAN ALGORITMA *N-GRAM* DAN ALGORITMA *BACKPROPAGATION*

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika

Oleh:

Serly Octalia
NIM : 09021381419074

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2018

LEMBAR PENGESAHAN TUGAS AKHIR

PENDETEKSIAN *MALWARE* ANDROID MENGGUNAKAN PENGGABUNGAN ALGORITMA *N-GRAM* DAN ALGORITMA *BACKPROPAGATION*

Oleh :

Serly Octalia
NIM : 09021381419074

Palembang, 31 Agustus 2018

Pembimbing I

Yoppy Sazaki, M.T.
NIPUS. 197406062015109101

Pembimbing II,

Kanda Januar Miraswan, M.T.
NIK. 1671080901900006

Mengetahui,
Ketua Jurusan Teknik Informatika

Rifkie Primartha, M.T.
NIP. 197706012009121004

TANDA LULUS UJIAN SIDANG TUGAS AKHIR

Pada hari Jum'at tanggal 31 Agustus 2018 telah dilaksanakan ujian sidang tugas akhir oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : **Serly Octalia**
NIM : **09021381419074**
Judul : **Pendeteksian *Malware* Android menggunakan Penggabungan Algoritma *N-Gram* dan Algoritma *Backpropagation***

1. **Pembimbing I**

   Yoppy Sazaki, M.T.
   NIPUS. 1671140201820005 .....

2. **Pembimbing II**

   Kanda Januar Miraswan, M.T.
   NIK. 1671080901900006 .....

3. **Penguji II**

   Rifkie Primartha, M. T.
   NIP. 197706012009121004 .....

4. **Penguji II**

   Mastura Diana Marieska, S.T., M. T.
   NIP. 198603212018032001 .....

Mengetahui,
Ketua Jurusan Teknik Informatika

Rifkie Primartha, M. T.
NIP. 197706012009121004

HALAMAN PERNYATAAN

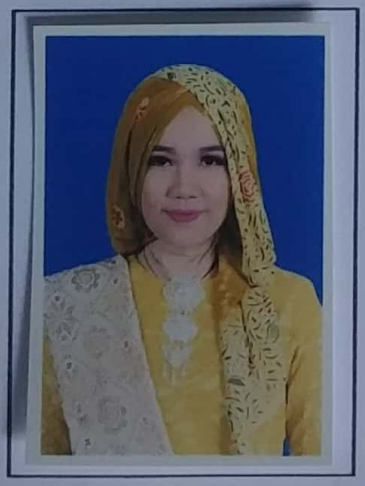
Yang bertanda tangan di bawah ini :

Nama	:	Serly Octalia
NIM	:	09021381419074
Program Studi	:	Teknik Informatika (Bilingual)
Judul Skripsi	:	Pendeteksian *Malware* Android menggunakan Penggabungan Algoritma *N-Gram* dan Algoritma *Backpropagation*

Hasil Pengecekan Software *iThenticate/Turnitin* : 2 %

Menyatakan bahwa Laporan Projek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan projek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 31 Agustus 2018

Serly Octalia
NIM. 09021381419074

Motto:

- Kerjakan apa yang bisa dikerjakan sekarang.
- Setiap pekerjaan dapat diselesaikan dengan mudah bila dikerjakan tanpa keengganan.
- Smile more, Talk less, Do more.
- Semangat, Sabar, dan Berdo'a.

Kudedikasikan karya tulis ini kepada :

- Allah SWT
- Ayahanda dan Ibunda Tercinta,
  H. Syafruddin, S.Sos. & Hj. Yamuna
- Kedua Kakakku,
  Andriansyah, S.T. & Herryadhi, S.Kom.
- Kedua Pembimbingku
- Almamater, serta
- Sahabat-sahabatku

ANDROID *MALWARE* DETECTION USING COMBINATION OF *N-GRAM* ALGORITHM AND *BACKPROPAGATION* ALGORITHM

By :
Serly Octalia
09021381419074

ABSTRACT

Malicious software development has become a serious threat to the security of android system, especially a new *malicious* software. There are new malicious software cannot be detected by conventional antivirus. This study created a framework for building software to detect and measure the accuracy of new and existing malware detection. In this study using combination of n-gram algorithm and backpropagation algorithm. N-gram algorithm is done statically so it needs a backpropagation algorithm so that the detection can be done dynamically. The application of combination n-gram algorithm and backpropagation algorithm successfully detect *Droiddream* type malware and this research can also detect malware variants, among other *DroidKungfu*, *BaseBridge* and *DroidCoupon* although it produce low accuracy. Therefore, this research can help in detecting a *Droiddream*, *DroidKungFu*, *BaseBridge* and *DroidCoupon* types of android malware on user devices for the safety and comfort of the user.

Keywords : Malicious software, Malware scanner, N-Gram, Backpropagation

Palembang, 31 Agustus 2018

Supervisor I

Yoppy Sazaki, M.T.
NIPUS.197406062015109101

Supervisor II,

Kanda Januar Miraswan, M.T.
NIK. 1671080901900006

Approve,
Chairman of Informatics Engineering Department

Rikie Primartha, M.T.
NIP. 197706012009121004

PENDETEKSIAN MALWARE ANDROID MENGGUNAKAN PENGGABUNGAN ALGORITMA N-GRAM DAN ALGORITMA BACKPROPAGATION

Oleh :
Serly Octalia
09021381419074

ABSTRAKSI

Perkembangan *malicious software* android menjadi ancaman yang cukup serius bagi keamanan sistem android, khususnya *malicious software* baru. Banyak *malicious software* baru tersebut tidak dapat dideteksi oleh sistem *antivirus* konvensional. Penelitian ini membuat sebuah kerangka kerja untuk membangun perangkat lunak dalam mendeteksi dan mengukur akurasi pendeteksian *malicious software* baru dan yang telah ada sebelumnya. Pada penelitian ini menggunakan penggabungan algoritma *n-gram* dan algoritma *backpropagation*. Algoritma *n-gram* dilakukan secara statis sehingga dibutuhkan algoritma *backpropagation* agar pendeteksian dapat dilakukan secara dinamis. Penerapan penggabungan algoritma *n-gram* dan algoritma *backpropagation* berhasil mendeteksi *malware* jenis *Droiddream* dengan baik dan penelitian ini juga dapat mendeteksi varian *malware* baru yaitu *DroidKungFu*, *BaseBridge* dan *DroidCoupon* walaupun menghasilkan akurasi yang rendah. Oleh karena itu, penelitian ini dapat membantu dalam mendeteksi *malware* android jenis *Droiddream, DroidKungFu, BaseBridge* dan *DroidCoupon* pada perangkat pengguna demi keamanan dan kenyamanan pengguna.

Kata Kunci : *Malicious software, Malware scanner, N-Gram, Backpropagation*

Palembang, 31 Agustus 2018

Pembimbing I

Yoppy Sazaki, M.T.
NIPUS.197406062015109101

Pembimbing II,

Kanda Januar Miraswan, M.T.
NIK. 1671080901900006

Mengetahui,
Ketua Jurusan Teknik Informatika

Rifkie Primartha, M.T.
NIP. 197706012009121004

KATA PENGANTAR

Alhamdulillāh, segala puji bagi Allah SWT semesta alam. Penulis memuji-Nya, memohon pertolongan dan ampunan-Nya, serta meminta perlindungan kepada-Nya dari kejahatan jiwa dan keburukan amal perbuatan yang telah dilakukan. Barangsiapa yang Allah beri petunjuk, tak seorangpun yang dapat menyesatkannya. Dan barang siapa yang Allah sesatkan, tiada seorangpun yang dapat memberinya petunjuk. Penulis bersaksi tiada Tuhan yang berhak disembah selain Allah dan Muhammad adalah hamba dan Rasul Allah.

Tiada ucapan terindah melainkan ucapan rasa syukur atas selesainya penyusunan tugas akhir yang berjudul "Pendeteksian *Malware* Android Menggunakan Penggabungan Algoritma *N-Gram* dan Algoritma *Backpropagation*". Tugas akhir tersebut merupakan syarat untuk memenuhi kurikulum Program Studi Teknik Informatika Universitas Sriwijaya.

Dalam penyusunan tugas akhir ini, penulis tidak terlepas dari hambatan dan tantangan. Namun demikian, berkat bimbingan, bantuan, dan dukungan dari berbagai pihak, akhirnya tulisan sederhana ini selesai tepat waktu. Ucapan terima kasih secara khusus penulis sampaikan kepada Ayah dan Ibu atas bantuan, dukungan baik materi maupun moril serta do'a yang selalu dipanjatkan untuk kebaikan dan keselamatan penulis di dunia dan di akhirat. Semoga tulisan sederhana ini dapat membuat Ayah dan Ibu bahagia. Tanpa mengurangi rasa hormat, penulis juga menyampaikan terima kasih kepada Bapak Yoppy Sazaki, M.T., Bapak Kanda Januar Miraswan, M.T. dan Dr. Mazura binti Mat Din selaku dosen pembimbing atas kesediaannya meluangkan waktu untuk membimbing penulis dari awal sampai akhir. Semoga amal kebaikan yang telah Bapak dan Ibu berikan bernilai pahala disisi Allah. Untaian kata di atas dirasa kurang indah jika tidak mengucapkan terima kasih kepada Bapak Rifkie Primartha, M.T. dan Ibu Mastura Diana Marieska, S.T, M.T. selaku dosen penguji atas kesediaannya memperbaiki kesalahan-kesalahan dalam penyusunan tugas akhir penulis.

Selain itu, penulis menyampaikan penghargaan dan terima kasih sebesar-besarnya kepada pihak-pihak yang telah membantu dalam penyelesaian tugas akhir ini, yaitu :

1. Kedua kakak penulis atas do'a, bantuan, dan dorongan semangat kepada penulis dalam menghadapi segala hal. Khususnya Saudaraku Herryadhi, S.Kom. yang selalu membantu penulis dalam membangun dan menguji *malware scanner* serta menganalisis *malware* pada penelitian ini.
2. Bapak Jaidan Jauhari, M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.

3. Bapak Rifkie Primartha, M.T., selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Alfarissi, M.Comp.Sc., selaku Pembimbing Akademik yang telah memberikan saran dan masukan selama perkuliahan.
5. Seluruh dosen dan staff administrasi Program Studi Teknik Informatika Universitas Sriwijaya atas bantuannya dalam perkuliahan dan penyusunan skripsi penulis.
6. Teman-teman seperjuangan Teknik Informatika Bilingual angkatan 2014, untuk persahabatan dan masa-masa perkuliahan yang menyenangkan dan tak terlupakan.
7. Kakak-kakak tingkat Teknik Informatika Bilingual angkatan 2013, khususnya Kak Ades Yudhatama, S.Kom. yang selalu memberikan bantuan baik dalam akademik maupun non-akademik serta saran kepada penulis.
8. Serta pihak-pihak lainnya yang terlibat selama pelaksanaan Tugas Akhir ini yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa masih banyak kekurangan dalam tugas akhir ini, untuk itu penulis mengharapkan saran dan kritik yang membangun demi penyempurnaan ke depan. Akhir kata, penulis berharap semoga tulisan ini dapat bermanfaat bagi semua pihak yang membutuhkan.

Palembang, 31 Agustus 2018

Serly Octalia
09021381419074

DAFTAR ISI

Halaman

DAFTAR TABEL

Halaman

DAFTAR GAMBAR

Halaman

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini membahas latar belakang masalah, rumusan masalah, tujuan, manfaat penelitian, dan batasan masalah. Bab ini akan memberikan penjelasan umum mengenai keseluruhan penelitian.

Pendahuluan dimulai dengan penjelasan mengenai proses pendeteksian *malware* android dengan penggabungan algorima *n-gram* dan algoritma *backpopagation* yang menjadi latar belakang dari penelitian ini.

1.2 Latar Belakang

Perkembangan teknologi seluler berjalan sangat cepat karena sejak awal tahun 1970 diterapkannya teknologi seluler yang hanya digunakan sebagai alat untuk komunikasi telefon dan memiliki ukuran yang cukup besar dan panjang, terus berkembang dari tahun ke tahun hingga sekarang menjadi teknologi seluler yang berukuran lebih kecil dan memiliki *operating system* yang dikenal dengan *smartphone*, sebab kecanggihan fitur yang dahulunya hanya ada di komputer, sekarang hampir semua dapat dilakukan oleh *smartphone* (Kompas Online, 2010). Salah satu *smartphone* yang popular saat ini adalah *smartphone* android karena berdasarkan data dari *International Data Center* (IDC) untuk kuartal 1 tahun 2017, pemasaran sistem operasi android mencapai 85% di seluruh dunia (www.idc.com). Dengan banyaknya pengguna android, toko-toko penyedia aplikasi android pun banyak bermunculan, begitu juga dengan ancaman keamanan yang semakin besar.

Ancaman yang serius saat ini adalah *malicious software* atau biasa disebut dengan *malware,* yang merupakan perangkat lunak memiliki kemampuan untuk mencuri data dari pengguna, merusak sistem, serta mengganggu kinerja sistem dengan melakukan banyak layanan di balik layar (Al Huda, F., *et al,* 2016). Ancaman *malware* dapat diatasi dengan mengembangkan sebuah *malware scanner,* secara sederhana didefinisikan sebagai sebuah program mencari sektor-sektor media penyimpanan dan file-file untuk rangkaian *byte* spesifik *malware* yang diketahui (Gryaznov,1999). Rangkaian *byte* ini disebut dengan *malware signature* karena metode pendeteksian dilakukan dengan pencocokan rangkaian *byte* spesifik *malware* android, pencarian dapat dilakukan dengan cepat dan tepat akan tetapi menghasilkan *false positives* dan *false negatives* yang tinggi serta memiliki kelemahan utama yaitu dibutuhkan *signature* tersendiri untuk dapat mendeteksi satu varian *malware*.

Penelitian mengenai pendeteksian *malware* telah banyak dilakukan, diantaranya adalah penelitian yang dilakukan oleh Park pada tahun 2006 yaitu menerapkan algoritma *n-gram*. Algoritma *n-gram* yang diterapkan pada pendeteksian *malware* bertujuan untuk menetapkan apakah file mengandung sebuah *malware* atau tidak.

Penelitian akan dikembangkan menjadi sebuah model baru yaitu dengan penerapan penggabungan *n-gram* dengan *backpropagation*. Penerapan algoritma *backpropagation* dan penarikan kesimpulan berdasarkan bobot dan kombinasi tingkah dapat memberikan kemampuan bernalar sehingga diharapkan pendeteksian *malware* baru maupun yang telah ada dapat segera dideteksi keberadaannya.

1.3 Perumusan Masalah

Perumusan masalah yang diangkat dalam penelitian ini adalah bagaimana membangun *malware scanner* untuk mendeteksi dan mengukur akurasi pendeteksian *malware* android menggunakan penggabungan algoritma *n-gram* dan algoritma *backpropagation*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk membangun *malware scanner* dan mengukur akurasi pendeteksian *malware* android dengan menerapkan penggabungan algoritma *n-gram* dan algoritma *backpropagation*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah membantu dalam mendeteksi keberadaan varian *malware* android pada perangkat pengguna demi keamaanan dan kenyamanan *user*.

1.6 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Lingkup pendeteksian *malware* terbatas hanya pada *malware* jenis *trojan* android yang diperoleh dari situs *http://virusshare.com/*;
2. Penelitian ini hanya dibatasi pada bagaimana cara pendeteksian *malware* android, tidak termasuk pembersihan dan *recovery* serangan *malware* android;

3. Aplikasi *malware scanner* yang dikembangkan yaitu berbasis android dan pelatihan data berbasis *desktop* tanpa adanya fasilitas *malware realtime protection*;
4. Penelitian ini terbatas hanya pada implementasi hasil analisis kode biner pada algoritma *n-gram* tanpa memberikan penjelasan detail tentang teknik analisis konde biner tersebut.

1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir ini mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya yaitu sebagai berikut.

BAB I PENDAHULUAN

Bab ini akan membahas tentang latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah atau ruang lingkup, meodologi penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini membahas seluruh dasar-dasar teori yang digunakan mulai dari definisi sistem, informasi mengenai domain, dan semua yang digunakan pada tahapan analisis, perancangan, dan implementasi.

BAB III METODOLOGI PENELITIAN

Bab ini membahas mengenai tahap-tahap yang akan diterapkan pada penelitian. Setiap rencana dari tahapan penelitian dideskripsiskan secara rinci berdasarkan kerangka kerja. Dilanjutkan dengan peancangan manajemen proyek dalam pelaksanaan penelitian.

BAB IV PENGEMBANGAN PERANGKAT LUNAK

Bab ini membahas perancangan dan lingkungan implementasi, berupa analisis dari masalah yang dihadapi dalam penelitian serta perancangan perangkat lunak untuk pendeteksian *malware* android menggunakan penggabungan algoritma *n-gram* dan algoritma *backpropagation* yang akan digunakan sebagai alat penelitian.

BAB V HASIL DAN ANALISIS PENELITIAN

Bab ini membahas implementasi dari hasil analisis dan perancangan yang sudah dilakukan sebelumnya. Hasil analisis berupa kesimpulan yang dapat diambil dari penelitian. Melakukan pengujian perangkat lunak dan pengujian data penelitian.

BAB VI KESIMPULAN DAN SARAN

Bab ini berisi semua kesimpulan dari uraian-uraian yang telah dibahas sebelumnya, dan saran yang diharapkan dapat berguna untuk pengembangan penelitian lebih lanjut.

1.8 Kesimpulan

Terdapat enam bab yang dibahas dalam penelitian ini. Bab 1 membahas gagasan dasar yang diajukan mengenai pendeteksian *malware* android menggunakan penggabungan algoritma *n-gram* dan algoritma *backpropagation*. Hal ini penting untuk memahami konsep dasar tentang apa yang akan dikerjakan. Latar belakang masalah dibahas pada bagian 1.2. Pernyataan masalah telah dijelaskan sehingga solusi untuk memecahkan masalah dapat diidentifikasi. Selain itu ada tujuan dalam penelitian dan ruang lingkup penelitian ini juga diberikan

dengan jelas. Pada bagian akhir, berisi alasan mengapa penting untuk melakukan penelitian ini.

Bab 2 akan membahas tinjauan dan kajian literature yang berkaitan dengan penelitian. Bab 3 adalah keseluruhan metodologi penelitian untuk merinci kerangka penelitian untuk mengembangkan sistem. Bab 4 akan melanjutkan tahap dari penelitian ini yaitu memberikan gambaran bagaimana perangkat lunak dikembangkan sehingga dapat menghasilkan hasil yang akurat beserta analisisnya seperti yang akan dituliskan pada Bab 5. Pada bagian akhir bab yaitu bab 6 yang menjadi kesimpulan dari penelitian yang diajukan, juga menerima saran dari semua pihak terkait penelitian ini agar dapat di perbaiki dan dikembangkan lebih lanjut.

DAFTAR PUSTAKA

Al Huda, F., Mahmudy, W. F., & Tolle, H. (2016). Android Malware Detection Using Backpropagation Neural Network. *Indonesian Journal of Electrical Engineering and Computer Science*, *4*(1), 240-244.

Alam, S., Qu, Z., Riley, R., Chen, Y., & Rastogi, V. (2017). DroidNative: Automating and optimizing detection of Android native code malware variants. *computers & security*, *65*, 230-246.

Fausett, L. 1993. Fundamentals of Neural Networks: Architectures, Algorithms And Applications. Prentice Hall, New Jersey, United States.

Gryaznov, D. 1999. Scanners of The Year 2000: Heuristics. Proceedings of the Fifth International Virus Bulletin Conference, pp.225-234.

Kiss, N., Lalande, J. F., Leslous, M., & Tong, V. V. T. (2016, May). Kharon dataset: Android malware under a microscope. In *The Learning from Authoritative Security Experiment Results (LASER) workshop* (pp. 1-12). USENIX Association.

Liu, Y., Zhang, Y., Wang, H., Xu, J., & Li, J. (2016, September). Research on standardization of the Android malware detection results. In *Network Infrastructure and Digital Content (IC-NIDC), 2016 IEEE International Conference on* (pp. 161-165). IEEE.

Palumbo, P., Sayfullina, L., Komashinskiy, D., Eirola, E., & Karhunen, J. (2017). A pragmatic android malware detection procedure. *Computers & Security*, *70*, 689-701.

Park, J. D. 2006. Cooperative Heuristics. Proceedings of the 16th Virus Bulletin International Conference 1(1):6-9.

Pratomo, B. A., & Ijtihadie, R. M. (2016). SISTEM DETEKSI INTRUSI MENGGUNAKAN N-GRAM DAN COSINE SIMILARITY. *JUTI: Jurnal Ilmiah Teknologi Informasi*, *14*(1), 108-116.

Purnomo, M. H. dan Kurniawan, A. 2006. Supervised Neural Networks dan Aplikasinya. Graha Ilmu, Yogyakarta, Indonesia.

Schultz, M. G. et al. 2001. Data Mining Methods for Detection of New Malicious Executables. Journal of IEEE Symposium on Security and Privacy 1(1):38-49, Oakland, CA, United States.

Tong, F., & Yan, Z. (2017). A hybrid approach of mobile malware detection in Android. *Journal of Parallel and Distributed Computing*, *103*, 22-31.

Wang, C., Li, Z., Gong, L., Mo, X., Yang, H., & Zhao, Y. (2017). An Android Malicious Code Detection Method Based on Improved DCA Algorithm. *Entropy*, *19*(2), 65.

Wang, X., Yang, Y., & Zeng, Y. (2015). Accurate mobile malware detection and classification in the cloud. *SpringerPlus*, *4*(1), 583.

Zheng, M., Sun, M., & Lui, J. C. (2013, July). Droid analytics: a signature based analytic system to collect, extract, analyze and associate android malware. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2013 12th IEEE International Conference on* (pp. 163-171). IEEE.

Zhu, H. J., You, Z. H., Zhu, Z. X., Shi, W. L., Chen, X., & Cheng, L. (2018). DroidDet: Effective and robust detection of android malware using static analysis along with rotation forest model. *Neurocomputing*, *272*, 638-646.