Teknik Pengesanan Botnet P2P Menggunakan Teknik Pemilihan Ciri

Mohammad Hairy Kharauddin
Kolej Komuniti Masjid Tanah
airis1482@yahoo.com

Wan Ahmad Ramzi Wan Yusuf
Kolej Komuniti Masjid Tanah
ramzi016@gmail.com

Mohd Fitri Ab Rasid
Kolej Komuniti Masjid Tanah
feetittho@yahoo.com

Abstract
Computer networks play an important role in modern society today. A wide range of business, communications, utilities, infrastructure, banking and leisure services are now provided by systems that rely on secure and efficient network operations. As network technology evolves rapidly in size and complexity, the necessary action is to understand the nature of the nature of the network to protect users from dangerous security threats. Botnets are identified as one of the most popular threats of their emergence and P2P botnets use P2P technology for file transfer techniques making these botnets difficult to detect and over the past decade researchers have given more focus in developing more efficient and effective botnet detection methods. In this research, supervised machine learning methods have been used and the main focus is on hybrid models using feature selection techniques or 'feature selection'. The accuracy and detection of botnets can be improved with the use of hybrid models which show an increase in accuracy from 69.89% to 87.82% detection of botnets in hybrid models compared to filter models. This study will help in identifying the characteristics of botnets and identify the characteristics of P2P botnets with the application of feature selection techniques with hybrid models.

Keywords: botnets, malware, feature selection

Abstrak
Rangkaian komputer memainkan peranan penting dalam masyarakat moden pada masa kini. Pelbagai jenis perniagaan, komunikasi, utiliti, infrastruktur, perbankan dan perkhidmatan santai kini disediakan oleh sistem yang bergantung pada operasi jaringan yang selamat dan efisien. Oleh kerana teknologi rangkaian berkembang secara pesat dalam ukuran dan kerumitan, tindakan perlu adalah untuk memahami sifat sifat rangkaian untuk melindungi pengguna dari ancaman keselamatan berbahaya. Botnet dikenal pasti sebagai salah satu ancaman yang paling popular kemunculannya dan botnet P2P menggunakan teknologi P2P untuk teknik tukaran fail menjadikan botnet ini sukar dikesan dan selama sedekad yang lalu para penyelidik memberikan lebih penumpuan dalam mengembangkan kaedah penyediaan pengesanan botnet yang lebih cekap dan berkesan. Dalam penyelidikan ini, kaedah pembelajaran mesin yang diselia telah digunakan dan tumpuan utama adalah pada model hybrid menggunapakai teknik

pemilihan ciri ataupun *'feature selection'*. Ketepatan dan pengesanan botnet ternyata dapat ditingkatkan dengan penggunaan model hybrid yang mana menunjukkan peningkatan ketepatan dari 69.89% hingga 87.82% pengesanan botnet pada model hibrid berbanding model penapis. Kajian ini akan membantu dalam mengenal pasti karektor botnet dan mengenal pasti ciri ciri botnet P2P dengan penerapan teknik pemilihan ciri dengan model hybrid.

Kata Kunci: botnet, perisian malware, pemilihan ciri

1.0. Pengenalan

Pada masa kini, rangkaian komputer memainkan peranan penting dalam masyarakat moden kita. Pelbagai jenis perniagaan, komunikasi, utiliti, infrastruktur, perbankan dan perkhidmatan rekreasi kini disediakan oleh sistem yang bergantung pada operasi jaringan yang aman dan efisien. Oleh kerana rangkaian berkembang secara besar-besaran dalam ukuran dan kerumitan, tindakan perlu adalah untuk memahami perilaku rangkaian untuk melindungi mereka dari ancaman berbahaya keselamatan. Ancaman keselamatan utama dan pembawa utama aktiviti jahat berasal dari Internet itu sendiri yang juga dikenali sebagai perisian malware. Trend terbaru dari malware adalah malware botnet yang terkenal dan malware botnet terdiri daripada banyak jenis virus seperti Trojan, rookits, worm dan banyak lagi (Kaspersky, 2013).

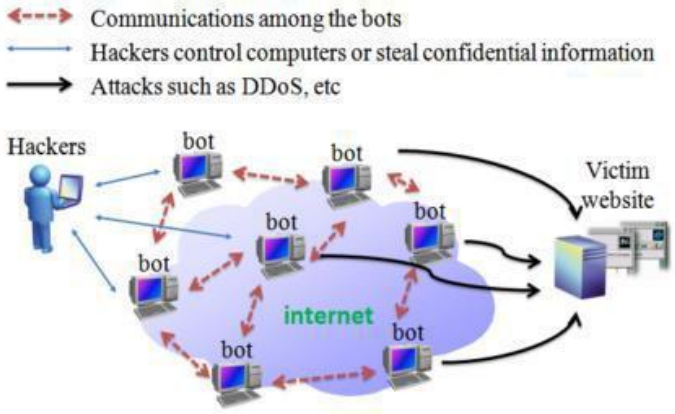
Botnet baru-baru ini dikenal pasti sebagai salah satu ancaman yang paling muncul dan botnet P2P menggunakan teknologi P2P untuk menukar fail menjadikan botnet jauh lebih sukar untuk dikesan. Botnet dapat dikenali dan dikesan dengan menganalisis perubahan ciri yang tidak normal dalam tingkah laku lalu lintas rangkaian. Makalah ini bertujuan untuk menyelidiki ciri botnet P2P dalam skala rangkaian tertentu dengan menganalisis lalu lintas dan tingkah laku rangkaian P2P yang meletakkan asasnya menggunakan alat analisis rangkaian yang berkesan. Analisis rangkaian akan membantu mengesan dan melakukan aktiviti yang tidak normal tepat pada masanya. Analisis yang dilakukan adalah sebahagian dari aktiviti projek dalam mengesan botnet di P2P.

P2 didefinisikan sebagai model rangkaian alternatif yang menyediakan seni bina pelayan pelanggan tradisional dan menggunakan model desentralisasi di setiap mesin dengan disebut sebagai *peer* dan fungsi yang disebut sebagai klien dalam lapisan fungsi pelayannya sendiri. Rakan sebaya berfungsi sebagai pelanggan dan pelayan pada masa yang sama (Vu, Lupu & Ooi, 2010). Rakan sebaya dapat memulai permintaan awal kepada kawanan lain sambil menanggapi permintaan masuk dari teman lain yang ada di jaringan. Rakan sebaya dapat mengatur diri mereka ke dalam kumpulan ad-hoc ketika mereka bekerjasama, berkomunikasi dan berkongsi lebar jalur antara satu sama lain dalam menyelesaikan tugas (Simon, 1991).

Botnet atau Bot Networks, dibina dari komputer jahat yang telah dijangkiti kod jahat yang dapat dikawal dari jarak jauh menggunakan perintah yang dihantar melalui internet (Ave, 2008). Komputer yang telah dijangkiti botnet disebut sebagai komputer 'zombie' yang memungkinkan

penggodam mengendalikan banyak komputer pada masa yang sama yang akan mengakibatkan penyebaran virus, menjana spam dan banyak lagi jenayah siber (Norton, 2015). Botnet P2P adalah salah satu ancaman keselamatan paling tinggi dan dikatakan lebih radikal dan lebih sukar untuk dikesan dan mencegah dibandingkan dengan bot lain. Dalam botnet P2P, bot dihubungkan dengan bot lain untuk menukar lalu lintas C&C, menghilangkan keperluan untuk pelayan terpusat yang mengakibatkan gangguan botnet P2P sukar untuk dianggarkan kerana ia sering menggunakan protokol khusus (Rossow & Andriesse, 2013).

Rajah 1 menunjukkan operasi Botnet P2P menyiratkan bahawa setiap mesin yang terinfeksi dalam kawanan akan bertindak sebagai rakan kawanan bagi yang lain dan oleh kerana itu kajian ini akan menggunakan pengesanan anomali yang melibatkan pemeriksaan penggunaan normal dan melaporkan apa yang tidak normal. Dengan beberapa parameter yang digunakan dalam teknik pengamatan dan pengelompokan, aliran normal P2P dan aliran abnormal botnet P2P akan diperhatikan dalam mengenali botnet P2P (Liao & Chang, 2010).



Rajah 1.0: Pengoperasian botnet P2P (Liao & Chang, 2010)

2.0 Kajian literatur

Terdapat banyak kaedah yang digunakan untuk mengesan botnet dan pembelajaran mesin adalah salah satu pilihan dalam mengesan botnet di rangkaian kami. Andaian utama kaedah berasaskan pembelajaran mesin adalah botnet akan membuatnya mempunyai sifat sendiri di dalam rangkaian dan sifat ini akan digunakan untuk mengesan botnet dengan berkesan menggunakan algoritma pembelajaran mesin (MLA). MLA memberikan pengesanan fleksibiliti dan memerlukan masa yang lebih singkat dalam mengesan botnet dalam rangkaian. Penyelesaian pengesanan pertama adalah dengan banyak sistem eksperimen telah dilaporkan dalam banyak penyelidikan yang telah dilakukan pada awal tahun 2000 dengan pelbagai tujuan dan berdasarkan pelbagai prinsip teknikal dan andaian tingkah laku botnet dan corak lalu lintas (Tyagi & Aghila, 2011).

Pengesanan botnet berasaskan rangkaian pertama menggunakan kaedah MLA digunakan oleh Livadas pada tahun 2006. Dia mencadangkan

pendekatan untuk menilai penggunaan MLA untuk mengenal pasti pendekatan yang dicadangkan dari botnet berbasis IRC. Dengan menggunakan tiga teknik MLA yang diselia untuk fasa klasifikasi iaitu pengkelasan Naïve Bayes, pengkelasan pohon keputusan dan algoritma C4.5 yang bertujuan untuk menunjukkan kecekapan teknik pembelajaran mesin yang berbeza dalam mengenal pasti lalu lintas botnet yang akan dinilai dengan teknik klasifikasi yang bervariasi, satu set pencirian atribut dan ukuran set latihan. Namun, penilaian yang dilakukan hanya memperoleh satu sampel malware botnet kerana botnet tersebut diperoleh hanya dari lalu lintas latar belakang kampus yang menjadikan sasaran hanya pada botnet berbasis IRC dan keberkesanannya dalam pelaksanaan dunia nyata sangat terbatas. Lebih-lebih lagi, kaedah ini rentan pada penghindaran oleh gangguan aliran.

Strayer et al. (2008), telah mendekati kaedah pengesanan berdasarkan tingkah laku rangkaian dan pembelajaran mesin dengan dan perluasan dan kerja yang dilakukan oleh Livadas. Sangat mirip dengan kaedah pendekatan Livadas, kerangka pengesanan botnet menggunakan beberapa pendekatan MLA untuk mengklasifikasikan aliran lalu lintas IRC sebagai botnet atau tidak hanya dengan tambahan pada penyelidikan Livadas sebelumnya. Menggunakan klasifikasi aliran yang diawasi yang sama, empat pengklasifikasi digunakan, C4.5, pohon keputusan, Naïve Bayes dan penambahan pengelasan rangkaian Bayesian. Sama dengan kaedah Livadas, penilaian Strayer kajian itu sendiri dilakukan pada eksperimen testbed yang dikendalikan sepenuhnya.

Kecekapan MLA dinilai menggunakan kadar positif palsu (FPR) dan negatif palsu (FNR). Kajian itu lebih sesuai dengan kemampuannya untuk mengesan botnet IRC dengan topologi terpusat dan ia memerlukan pertimbangan tambahan baik oleh manusia atau mesin untuk mengetahui keberadaan botnet. Kekurangan dari pendekatan ini adalah hanya dapat memodelkan pola lalu lintas TCP sebagai pembawa utama komunikasi IRC.

Nogueira, Salvador & Blessa, (2010), telah mengusulkan pendekatan pengesanan botnet menggunakan teknik pembelajaran mesin yang memberikan pengenalan lalu lintas botnet menggunakan Artificial Neural Networks (ANNs) sebagai klasifikasi algoritma. Perbezaan pendekatan ini adalah ia memiliki kemampuan untuk beroperasi secara on-line dan mudah menyesuaikan diri dengan setiap perubahan dalam pola lalu lintas botnet sehingga menciptakan fitur baru untuk kerangka umum pengesanan botnet menggunakan teknik pembelajaran mesin yang diawasi. Entiti tambahan yang juga dikenali sebagai Intrusion Management System (IMS) yang menggunakan pemerhatian lalu lintas yang dilakukan oleh sistem pengesanan serta hasil klasifikasi untuk melihat hasil kehadiran lalu lintas berbahaya. Walau bagaimanapun, pendekatan ini masih mempunyai beberapa kelemahan contohnya keperluan untuk penilaian luaran agar disediakan fungsi penyesuaian.

Sistem pengesanan botnet berdasarkan aliran dalam mengenali lalu lintas botnet yang berbahaya dan sistem yang disasarkan pada aliran botnet IRC dan TCP dengan menggunakan pemeriksaan forensik dan paket

mendalam (DPI) host dalam mengekstrak ciri lalu lintas. Kaedah ini menggunakan lima pengklasifikasi berbeza iaitu Naïve Bayesian, jaringan Bayesian, Support Vector Machine (SVM), Boosted keputusan pohon dan C4.5 *tree tree*. Kekurangan utama teknik ini adalah penggunaan forensik tahap klien dan DPI untuk memperoleh ciri aliran.

Saad et al. (2011), mencadangkan sistem untuk mengesan botnet P2P dengan menggunakan pemerhatian dan analisis lalu lintas di mana ciri-ciri lalu lintas berdasarkan hos dan aliran diperhatikan. Penulis menggunakan lima MLA yang diselia iaitu Jiran terdekat, Bayesian naif, Mesin Vektor Sokongan (SVM), Rangkaian Neural tiruan (ANN) dan pengklasifikasi berasaskan Gauss. Hasilnya, SVM mencapai ketepatan 97%. Penulis bagaimanapun tidak memberikan penilaian mengenai pengelasan pokok yang mampu dan membuat lebih banyak kajian perlu dilakukan.

Walaupun semua penyelidikan yang dinyatakan di atas memberikan sistem persembahan pengesanan botnet yang cukup baik pada kumpulan data tertentu tertentu dari kaedah kontemporari tidak memberikan pandangan yang komprehensif mengenai Pohon Keputusan secara khusus untuk pengesanan pencerobohan dalam rangkaian menggunakan MLA, dan para penyelidik juga tidak menyiasat bagaimana banyak trafik diperlukan untuk diperhatikan setiap aliran supaya trafik botnet jahat dan tidak berbahaya dapat berjaya direkodkan.

Makalah ini bertujuan untuk mengatasi kekurangan ini melalui sistem pengesanan berdasarkan novel menggunakan pembelajaran MLA yang diselia menggunakan model filter dan hibrid khususnya Na specificallyve Bayes dan Naïve Bayes dengan Algoritma Genetik untuk mengenal pasti lalu lintas botnet.

3.0 Methodology

Metodologi adalah sistem prinsip atau peraturan yang luas dari kaedah atau prosedur tertentu yang dapat diturunkan untuk memahami situasi yang berbeza atau menyelesaikan masalah yang berbeza dalam ruang lingkup disiplin tertentu. Di bahagian ini, menerangkan pendekatan untuk mengesan kehadiran botnet P2P berdasarkan pengujian set data individu dengan kedua-dua teknik menggunakan pemilihan ciri hybrid model dan ciri penapis model pemilihan. Setiap set data individu dijangka dihantar hasil varian botnet dan teknik yang digunakan.

3.1 Naive Bayes Classifier

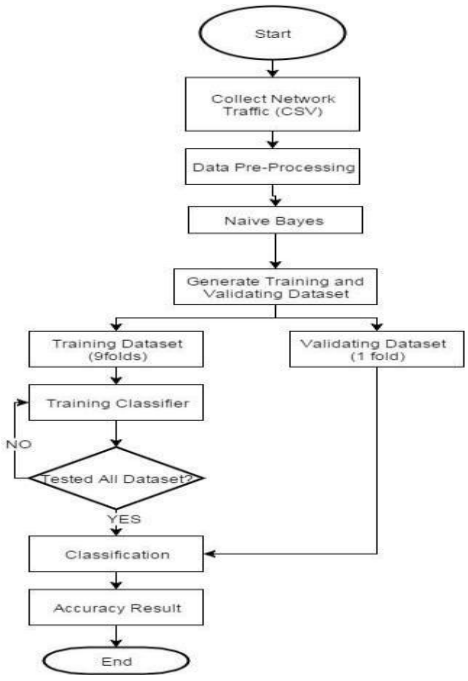
Klasifikasi Naïve Bayesian berfungsi dengan menganggap tidak ada ketergantungan antara atribut, yang juga dikenal sebagai kebebasan bersyarat kelas. Ia dilakukan untuk mempermudah pengiraan dan inilah sebabnya klasifikasi ini disebut 'naif'. Pengelaskan ini juga dikenali sebagai Bayes idiot, Bayes sederhana atau Bayes bebas (Miquélez et al. 2004). Naïve Bayes adalah teknik yang sangat intuitif di mana tidak seperti rangkaian saraf, teknik ini tidak mempunyai beberapa parameter bebas yang mesti ditetapkan yang sangat memudahkan proses reka bentuk. Klasifikasi Naïve Bayes mengembalikan kebarangkalian dan tidak memerlukan sejumlah besar data sebelum pembelajaran dapat dimulakan. Ia juga pantas ketika

membuat keputusan berbanding dengan jenis pengklasifikasi lain (Stern et al. 1999).

3.2 Filter Model

Algoritma Naive Bayes adalah kebarangkalian bersyarat yang menggunakan Teorema Bayes, formula yang mengira kebarangkalian dengan menghitung kekerapan nilai dan kombinasi nilai dalam set data. Dalam model saringan, data yang dikumpulkan melalui pra proses data di mana hingar data yang ada dalam kumpulan data dikurangkan dan dilatih oleh Naïve Bayes kemudian mengklasifikasikan melalui hasil pembelajaran data latihan. Setiap attribute bersyarat akan diberi label kelas untuk menguji ketepatan data set ujian. Set data akan diuji sebanyak 10 kali atau dinamakan 10 'fold' di mana 9 lipatan akan digunakan sebagai data latihan dan 1 lipatan akan digunakan sebagai mengesahkan set data. Peraturan Naïve Bayes adalah untuk mengira kebarangkalian kelas menggunakan contoh sifat dan ramalan kelas dilakukan dengan mengenal pasti kelas dengan kebarangkalian posterior tertinggi.

Hasil ketepatan akan dipaparkan selepas itu. Pengiraan menggunakan algoritma ini adalah dengan membuat andaian bahawa semua atribut adalah bebas bersyarat memandangkan nilai kelas. Naïve Bayes sebagai kaedah klasifikasi standard dalam pembelajaran mesin digunakan dengan baik kerana mudah diprogramkan, intuitif dan juga cepat melatih dan dapat menangani atribut yang hilang dengan mudah. Semua lapan set data akan dijalankan dengan pengelasan ini dan ketepatannya akan direkodkan untuk perbandingan kemudian.

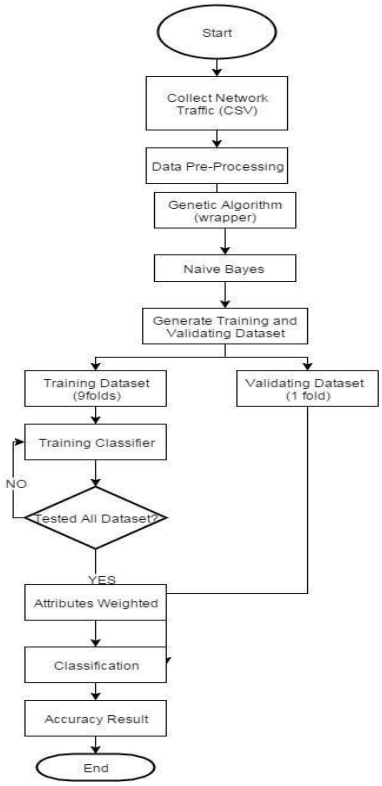


Rajah 2.0: Aliran Model Tapisan untuk Pendekatan Analisis Data

3.3 Hybrid Model

Set data telah melalui proses pra-pemprosesan yang serupa dengan Model Penapis. Model Hibrid menggunakan algoritma genetik berfungsi sebagai pembungkus untuk algoritma pemilihan ciri dan melakukan pencarian untuk subset yang betul menggunakan algoritma induksi itu sendiri sebagai bahagian fungsi penilaian. Pada langkah pertama, tujuannya adalah untuk mengurangi atribut yang ditetapkan dengan mengenal pasti subset atribut yang paling informatif untuk botnet menggunakan. Algoritma Genetik digunakan untuk mengenal pasti Naïve Bayes yang sebaliknya akan mengoptimumkan masalah seperti yang dinyatakan dalam pertandingan. Di sini parameter atribut optimum akan dipilih dan dilatih menggunakan Naïve Bayes sebanyak 10 kali ganda.

Subset ciri awal mengandungi semua atribut yang ada dalam set data. 9 lipatan akan digunakan untuk ujian dan 1 lipatan akan digunakan untuk mengesahkan set data latihan. Hasil ketepatan akan dipaparkan. Pengiraan menggunakan algoritma ini adalah mungkin dengan membuat andaian bahawa semua atribut bebas bersyarat memandangkan nilai kelas. Naïve Bayes sebagai kaedah klasifikasi standard dalam pembelajaran mesin digunakan dengan baik kerana mudah diprogramkan, intuitif dan cepat melatih dan dapat menangani atribut yang hilang dengan mudah. Semua lapan set data akan dijalankan dengan pengelasan ini dan ketepatannya akan digunakan untuk perbandingan dengan model penapis.



Rajah 3.0: Aliran Model Hibrid untuk Pendekatan Analisis Data

4.0 Hasil dan Perbincangan

Jadual 1 merupakan klasifikasi fitur botnet digunakan sebagai metrik pengukuran prestasi dalam menilai model pengesanan. Model yang baik akan menunjukkan kadar ketepatan pengesanan yang lebih tinggi yang baik untuk meramalkan tingkah laku tidak normal dalam lalu lintas rangkaian, sehingga berkesan dapat mengesan serangan botnet P2P.

Jadual 1: Fitur Botnet yang digunakan dalam kajian ini

port_a	sack_pkts_sent_b2a	actual_data_bytes_b2a
port_b	dsack_pkts_sent_a2b	rexmt_data_pkts_a2b
total_packets_a2b	dsack_pkts_sent_b2a	rexmt_data_pkts_b2a
total_packets_b2a	max_sack_blks/ack_a2b	rexmt_data_bytes_a2b
resets_sent_a2b	max_sack_blks/ack_b2a	rexmt_data_bytes_b2a
resets_sent_b2a	unique_bytes_sent_a2b	zwnd_probe_pkts_a2b
ack_pkts_sent_a2b	unique_bytes_sent_b2a	zwnd_probe_pkts_b2a
ack_pkts_sent_b2a	actual_data_pkts_a2b	zwnd_probe_bytes_a2b
pure_acks_sent_a2b	actual_data_pkts_b2a	zwnd_probe_bytes_b2a
pure_acks_sent_b2a	actual_data_bytes_a2b	outoforder_pkts_a2b
sack_pkts_sent_a2b	sack_pkts_sent_b2a	outoforder_pkts_b2a
req_1323_ts_a2b	avg_segm_size_b2a	pushed_data_pkts_a2b
req_1323_ts_b2a	max_win_adv_a2b	pushed_data_pkts_b2a
adv_wind_scale_a2b	max_win_adv_b2a	SYN _pkts_sent_a2b
adv_wind_scale_b2a	min_win_adv_a2b	SYN _pkts_sent_b2a
req_sack_a2b	min_win_adv_b2a	initial_window_bytes_b2a
req_sack_b2a	zero_win_adv_a2b	initial_window_pkts_a2b
sacks_sent_a2b	zero_win_adv_b2a	initial_window_pkts_b2a
sacks_sent_b2a	avg_win_adv_a2b	truncated_data_a2b
urgent_data_pkts_a2b	avg_win_adv_b2a	truncated_data_b2a
urgent_data_pkts_b2a	initial_window_bytes_a2b	truncated_packets_a2b
idletime_max_b2a	throughput_a2b	truncated_packets_b2a
hardware_dups_a2b	throughput_b2a	data_xmit_time_a2b
hardware_dups_b2a	class	data_xmit_time_b2a
FIN_pkts_sent_a2b	req_1323_ws _a2b	idletime_max_a2b
FIN_pkts_sent_b2a	req_1323_ws _b2a	

Dalam makalah ini, kami menjalankan teknik klasifikasi untuk membersihkan dan menormalkan set data yang menghasilkan 77 atribut dari 89 atribut dipilih dan diuji. Terdapat tujuh varian Botnet P2P yang telah dikenal pasti untuk digunakan dalam eksperimen ini seperti ditunjukkan dalam **Jadual 2**. Set data dari 7 varian botnet dibahas melalui dua teknik pemilihan ciri yang berbeza dengan wrapper dan tanpa wrapper menurut Model Filter dan Hybrid Model. Hasil kedua model yang diuji berdasarkan ketepatan akan dibandingkan dan dianalisis.

Jadual 2: Perincian Analisis Set Data

CSV Names	Botnet			Type of Traffic
Cryptowall.csv	Cryptowall	1.42MB	5088	4 = Abnormal 0 =Normal
Kelihos.csv	Kelihos	11.6MB	42450	2= Abnormal 0= Normal
Neris.csv	Neris	7.25MB	26592	8= Abnormal 0=Normal
Rbot.csv	Rbot	1.46MB	5237	7=Abnormal 0=Normal
Tbot.csv	Tbot	1.45MB	5179	3=Abnormal 0=Normal
Zbot.csv	Zbot	3.51MB	11644	6= Abnormal
				0=Normal
Zeroaccess.csv	Zeroaccess	1.43MB	5131	9 = Abnormal 0=Normal

5.0 Keputusan

Bahagian ini memaparkan semua hasil jadual klasifikasi untuk setiap model dalam model filter dan model hibrid dengan membandingkan antara semua varian botnet kumpulan data.

Jadual 3: Ketepatan untuk setiap model

Accuracy In Each Dataset			
	Naïve Bayes	Genetic Algorithm & Naïve Bayes	Increment of Accuracy
Crypto Wall	99.96%	99.98%	0.02%
Neris	91.44%	94.26%	2.85%
Kelihos	98.69%	98.76%	0.07%
Rbot	69.58%	99.92%	30.34%
Tbot	99.88%	99.92%	0.04%
Zbot	99.85%	99.96%	0.11%
ZeroAccess	99.98%	100.00%	0.02%
FullBotnet	69.89%	87.82%	17.93%

Jadual 4: Graf perbandingan ketepatan untuk pengesanan model Penapis dan Hibrid.

Terdapat lapan trafik rangkaian dengan label berlainan data. Terdapat trafik rangkaian biasa yang dilabel sebagai 0, CryptoWall botnet dilabel sebagai 4, Neris dilabel sebagai 8, Kelihos dilabelkan sebagai 2, Rbot dilabelkan sebagai 7, Tbot dilabelkan sebagai 3, Zbot dilabelkan sebagai 6, dan ZeroAccess dilabel sebagai 1. Model penapis digunakan Naïve Klasifikasi Bayes sementara model Hybrid menggunakan Algoritma Genetik (GA) dan Naïve Bayes.

Dalam setiap kumpulan data varian botnet, pengesanan botnet model hybrid mempunyai hasil yang lebih baik dibandingkan dengan model filter yang digunakan dalam mengesan botnet. Ketujuh varian botnet menunjukkan peningkatan dalam prestasi ketepatan ketika model hibrid digunakan sebagai kerangka pengesanan botnet dan bukan hanya

menggunakan model penapis. Ini bermaksud setiap varian botnet menunjukkan peningkatan pengesanan pada model hibrida dibandingkan dengan model penapis.

Untuk memperkukuhkan perbandingan kedua model ini, wakil varian baru dari setiap model yang dilabelkan sebagai FullBotnet telah dibuat yang terdiri dari semua 7 varian botnet. Dalam set data bonet penuh juga diuji dan hasilnya dalam **Jadual 3**. Manakala **Jadual 4** menunjukkan peningkatan ketepatan dari 69.89% hingga 87.82% pengesanan botnet pada model hibrid berbanding model penapis.

6.0 Kesimpulan

Ringkasnya, hasil pengujian disahkan menggunakan pengesahan silang 10 kali ganda. Dari hasil pengujian, terbukti bahawa dengan menggunakan model hibrid untuk pemilihan fitur, yang dalam projek ini adalah Genetik Algoritma sebagai pembungkus dan Naïve Bayes sebagai klasifikasi, ketepatan pengesanan botnet ditingkatkan. Oleh itu, hasilnya jelas menunjukkan bahawa model hibrid lebih baik dalam mengenali pengesanan dan ketepatan botnet P2P berbanding dengan model penapis.

Rujukan

Abdullah, R. & Abdollah, M., 2013. Revealing the Criterion on Botnet Detection Technique. *IJCSI International Journal of Computer Science Issues,* 10(2), pp.208–215.

Abdullah, R. & Ud, M.M., 2011. Recognizing P2P Botnets Characteristic Through TCP Distinctive Behaviour. *IJCSI International Journal of Computer Science Issues*, 9(12), pp.12–16.

Huawei, (2013). *Botnets and ddos attacks report.*

Kaspersky, (2013). What is a botnet? -kaspersky daily | kaspersky lab Retrived from http://blog.kaspersky.com/botnet/

Liao, W. & Chang, C.-C., (2010). Peer to peer botnet detection using data mining scheme. *In International Conference on Internet Technology and Applications*, pp.1–4.

Nogueira, A., Salvador, P. & Blessa, F., (2010). A botnet detection system based on neural networks. *Digital Telecommunications (ICDT), 2010 Fifth International Conference on*, pp.57–62.

Norton Symantec, Bots and Botnets—A Growing Threat. Retrieved from: http://us.norton.com/botnet/ [Accessed May 2, 2020].

Rossow, C. & Andriesse, D., (2013). *Sok: P2pwned-modeling and evaluating the resilience of peer-to-peer botnets*.

Simon, S., (1991). Peer-to-peer network management in an IBM SNA network. *IEEE Network*, *5*(2), pp.30–34.

Strayer, W. T., Walsh, R., Livadas, C. & Lapsley, D., (2006), Detecting botnets with tight command and control, *In Annual IEEE Conference on Local Computer Networks (LCN '06)* (pp. 195–202). Tampa, Florida, USA.

Tanenbaum, A., (2003). The analysis and identification of P2P botnet's Traffic Flows. *International Journal of Communication Networks & Information Security*, *3*(2),138–149.

Tyagi, A.K. & Aghila, G. (2011). A wide scale survey on botnet. *In International Journal of Computer Applications*, 34(9), 9–22.

Tyson, J.,(2000). *How the old napster worked – how stuff works*. Retrive from http://computer.howstuffworks.com/napster.htm.

Vu, Q.H., Lupu, M. & Ooi, B.C., 2010). *Peer-to-peer computing: principles and applications.*, pp.1– 317.

Wang, P., Sparks, S. & Cou, C., (2008). *Practical Machine Learning Tools and Techniques (Google eBook)*, Retrived from http://books.google.com/books?id=bDtLM8CODsQC&pgis=1.

Yeshwantrao, S. a & Jadhav, P.V.J., (2014). *Threats of Botnet to Internet Security and Respective Defense Strategies.*, *4*(1),121–127.