Studi Pengklasifikasi Berdasarkan Machine learning Untuk Mendeteksi Malware Berbasis Metode PE Probe

1st Aria Fajar Pratana
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
ariafajar@students.telkomuniversity.ac.id

2nd Satria Mandala
Fakultas Informatika
Universitas Telkom
Bandung, Indonesua
satriamandala@students.telkomuniversity.ac.id

Abstract*—Malware adalah perangkat lunak atau program berbahaya yang dapat merusak sistem operasi komputer dan jaringan. Banyak penelitian telah dilakukan untuk mencegah serangan malware yang dapat menimbulkan kerugian. Para peneliti telah mengembangkan metode deteksi dan klasifikasi malware berdasarkan metode statis dan dinamis. Namun, kedua metode ini memiliki kekurangan masing-masing, seperti metode statis yang kurang efektif dalam mendeteksi jenis file baru, dan metode dinamis yang lebih memakan sumber daya dan memiliki biaya yang lebih tinggi. Para peneliti juga telah mengembangkan berbagai teknik untuk mendeteksi malware berdasarkan fitur-fitur Pe-Probe. Untuk mengatasi masalah tersebut, tugas akhir ini mengusulkan pengembangan algoritma deteksi malware berbasis fitur Pe-Probe menggunakan machine learning untuk meningkatkan akurasi deteksi dan klasifikasi. Metode yang digunakan dalam penelitian ini adalah studi literatur tentang deteksi dan klasifikasi malware, pengembangan algoritma klasifikasi untuk deteksi malware berbasis metode Pe-Probe, pengembangan prototype, pengujian performansi, dan analisis. Penelitian ini menggunakan dataset sebanyak 19611 file Pe-Probe, yang terdiri dari data Pe-Probe yang terinfeksi malware dan tidak terinfeksi malware. Dengan menggunakan model machine learning Bagging, hasil menunjukkan nilai akurasi sebesar 0.976, detection rate sebesar 0.9881, dan false alarm rate sebesar 0.062.

Keywords—Malware, PE PROBE, Machine Learning.

I. INTRODUCTION

Dengan pesatnya perkembangan internet, malware telah menjadi salah satu ancaman utama di dunia cyber saat ini. Malware dapat berupa perangkat lunak berbahaya, pencurian data dan informasi, serta spionase. Menurut definisi Kaspersky Labs [7], malware adalah program komputer yang dirancang untuk menginfeksi komputer pengguna dan menimbulkan kerusakan di dalamnya dengan berbagai cara.

Perlindungan dari malware pada sistem komputer adalah tugas keamanan cyber yang penting bagi pengguna karena satu serangan saja dapat mengakibatkan kerusakan yang cukup besar. Saat ini, metode populer dalam mendeteksi malware menggunakan teknik klasifikasi statis dan dinamis yang berdasarkan algoritma yang mempelajari signature based dari malware, seperti yang dijelaskan oleh Chumachenko [3].

Namun, terdapat banyak pro dan kontra mengenai metode statis dan dinamis. Kelebihan teknik analisis statik adalah malware tidak dijalankan, sehingga mengurangi risiko terinfeksi malware. Meskipun begitu, metode ini memiliki keterbatasan, seperti tidak dapat mendeteksi jenis malware baru. Saat ini, penulis malware sering menggunakan teknik yang membuat proses analisis malware statik semakin sulit, seperti menggunakan packer untuk melakukan modifikasi kode secara otomatis. Di sisi lain, analisis dinamis lebih efisien dan tidak memerlukan executable untuk dibongkar atau didekripsi. Namun, metode ini memakan waktu dan sumber daya yang banyak, seperti yang dijelaskan oleh L. Nataraj dan S. Karthikeyan [11].

Penelitian lainnya, seperti yang dilakukan oleh Sungtaek OH, Woong Go, dan Taejin Lee [12], menggunakan deteksi malware berbasis signature-based dengan mengidentifikasi ciri-ciri malware yang berada di database. Namun, metode ini tidak sepenuhnya optimal ketika terjadi serangan zero-day, dan banyak malware yang tidak terdeteksi.

Pada tahun Vyas [20].melakukan penelitian dengan menginvestigasi fitur statis untuk mengusulkan sistem deteksi malware jaringan waktu nyata. Fitur-fitur tersebut diekstraksi dari header dan bagian file PE dan dikelompokkan menjadi empat kategori: metadata file, file pengepakan, DLL yang diimpor, dan fungsi yang diimpor. Fitur-fitur ini kemudian digunakan untuk melatih beberapa mesin pengklasifikasi pembelajaran.

Penelitian lainnya, seperti yang dilakukan oleh Tina Razei [15] menggunakan teknik deteksi malware berbasis Pe-Probe (Pe file) yang berdasarkan 9 fitur dari pe header dengan tingkat akurasi 95,5% menggunakan Random Forest.

Dari permasalahan itu, penelitian ini berfokus pada pengembangan deteksi dan klasifikasi akurasi malware berbasis feature-feature yang terdapat dalam pe-probe (pe file), dengan membuat sebuah prototype untuk deteksi malware dengan menggunakan machine learning untuk mendeteksi dan mengklasifikasikan akurasi dari PE-Probe (pe file). Diharapkan penelitian ini dapat menjadi metode optimal dalam mendeteksi malware.

II. RELATED WORK

Penelitian tentang deteksi malware sudah bermula sejak tahun 2005. Berikut adalah 15 penelitian terkait yang sudah dipublikasikan sejak tahun 2015 sampai sekarang.

Han [5] melakukan penelitian dengan metode 4-LFE dengan cara mengekstrak multi-fitur dari program jahat dengan menggabungkan fitur piksel dan n-gram untuk mengklasifikasi malware. setelah melaku kanpengujian kepada data publik terdiri dari 10.868 sampel malware, mencapai nilai akurasi sebesar 99,99%.

Dong Hee-Kim [6] melakukan penelitian deteksi malware menggunakan fitur PE-Header.algoritma yang digunakan dalam penelitian deteksi malware dengan cara menggabungkan algoritma Cart, SVC dan SGD. Hasil penelitian ini menghasilkan akurasi sebesar 99,99%

Liu [8] mengusulkan sistem analisis malware berbasis pembelajaran mesin, yang terdiri dari tiga modul pemrosesan data, pengambilan keputusan, dan deteksi malware baru. modul pemrosesan data menggunakan fitur gray-scale images, Opcode n-gram, dan import functions, untuk modul pengambilan keputusan menggunakan fitur klasifikasi dan mengindentifikasi malware, dan modul deteksi malware menggunakan fitur algoritma SNN. pengujian menggunakan 20.000 contoh malware dengan hasil akurasi sebesar 86,7%.

Udayakumar [19] mengusulkan teknik baru dalam mendeteksi dan klasifikasi malware dengan teknik analisis menggunakan alat Knime dan Orange, menggunakan 6 algoritma yang berbeda. dari hasil pengujian data set pada mongo DB algoritma Random Forest mencapai nilai akurasi sebesar 63,49% pada Knime dan 94,2% pada Orange

R.J.Mangialardo dan J.C.Duarte [10] melakukan percobaan dengan menggabungkan analisis statis dan dinamis dalam menganalisis klasifikasi dan indentifikasi malware. Pengujian ini menggunakan algoritma C.50 dan Random Forest yang di implementasikan dalam Framework. Percobaan ini menghasilkan akurasi sebesar 95,75%.

[13] melakukan penelitian klasifikasi malware dengan teknik clustering menggunakan algoritma K-means dan Expectation Maximizational. Metode clustering yang digunakan berdasarkan perhitungan skor Hidden Markov Model, memvariasikan jumlah cluster dari 2 dan 10, dan jumlah dimensi (skor) dari 2 hingga 5. pengujian ini menggunakan 8000 sampel malware dan menghasilkan nilai akurasi sebesar 90%

Shafiq [18] juga melakukan penelitian deteksi malware serangan Zero-day dalam struktur framework PE-Miner menggunakan fitur yang distandarisasi oleh sistem operasi Microsoft untuk file executable, DLLs dan file objek. Pengujian di lakukan selama 1jam dapat mengeskusi kumpulan data sebanyak lebih dari 17ribu data dan menghasilkan nilai akurasi sebesar 99% .

Xue [23] juga melakukan percobaan dalam mengklasifikasi malware berdasarkan penilaian probabilitas dan machine learning. Tahap 1 menggunakan neural networks dengan Spatial Pyramid Pooling untuk menganalisi gambar grayscale(fitur dinamis), Tahap 2 menggunakan variable n-gram dan machine learning, dan malscore digunakan untuk menggabungkan tahap 1 dan tahap 2. dari eksperimen pada 174.607 sampel malware dari 63 jenis malware, malscore dapat mengklasifikasi dengan nilai akurasi sebesar 98,82%.

Abijah Roselin [1] juga mengusulkan sebuah metode sistem mendeteksi malware pendekatan visualisasi 2d dengan digabungkan dengan pembelajaran mesin. (machine learning). Hasil dari pengujian sistem terhadap teknik yang lebih canggih seperti Malimg, BIG2015, dan MaleVis malware datasets, memiliki tingkat akurasi sebesar 98,65%, 97,2%, dan 97,43%.

Di tahun yang sama chen [2] melakukan percobaan deteksi dan klasifikasi malware dengan mengimplementasikan teknik klasifikasi menggunakan support vector machine (SVM) dengan menggabungkan algortima pembelajaran aktif (ALBL) untuk melakukan klasifikasi malware. pengujian dilakukan terhadap data malware yang sudah dikumpulkan oleh Microsoft Malware Classification Challenge (BIG 2015) di Kaggle dan mampu meningkatkan performa machine learning dalam mendeteksi dan klasifikasi malware.

Das [4] juga melakukan percobaan dalam mendeteksi dan klasifikasi malware secara online dengan meningkatkan perangkat keras, dan Guard OL (gabungan prosesor dan FPGA). Algortima yang digunakan adalah multilayer perceptron untuk melatih pendeteksian dini malware berbahaya atau jinak. metode tersebut menghasilkan nilai akurasi pada prediski awal sebesar 46% malware dalam 30% eksekusi pertama, sedangkan 97% sampel dari 100% setelah eksekusi

Mangialardo, R. J., and Duarte, J. C. [10] melakukan penelitian pengklasifikasian malware dengan cara memvisualisasikan malware. metode yang digunakan adalah LBP (Local Binary Patterns) dan (Scale-invariant feature transform) dengan mengelompokan malware kedalam blok menggunakan model bag-of-visual-words (BoVW). pengujian dilakukan terhadap 3 database malware dan hasil pengujian dapat meningkatkan klasifikasi yang cukup baik dan canggih.

Sahu [17] juga mengusulkan teknik klasifikasi dan deteksi malware berbasis grafik yang digunakan untuk mengumpulkan fitur malware yang berbeda. metode yang digunakan adalah metode hybrid, berdasarkan DAG dan Gaussian Support Vector Machines, untuk klasifikasi malware. pengujian dilakukan berdasarkan data KDD cup 1999 dan memberikan hasil akurasi yang tinggi dengan tingkat kesalahan pendeteksian di bawah 1%.

Wuchner [22] juga mengusulkan teknik deteksi malware menggunakan metode compression-based mining pada grafik aliran data kuantitatif. Dari hasil pengujian yang dilakukan terhadap kumpulan malware yang beragam, hasil pengujian mengungguli model deteksi berbasis frekuensi dalam hal efektifitas sebesar 600%.

Shiming Xia [23] melakukan penelitian deteksi malware berbasis gambar. metode yang digunakan menggunakan SVM untuk mendeteksi malware file, dan metode Markov Transition Field (MTF) untuk memeriksa dan klasifikasi malware yang nantinya akan di rubah kembali dalam gambar satudimensi ke bentuk vektor SVM. Penelitian

tersebut dilakukan didataset danmenghasilkan yang lebih baik di bandingkan metode berbasis gambar grayscale byteplot

III. Methodology and Dataset

A. *Dataset*

Data yang digunakan dalam melakukan penelitian ini adalah data malware yang berbentuk fitur ektrasi dalam pe-probe(pe-files) dengan jumlah data sebanyak 19610 jenis pe file yang mengandung malware dan tidak mengandung malware.

Table 1: Feature dataset pe probe yang digunakan machine learning

Header	Feature
Dos Header	e_magic, e_cblp, e_cp, e_crlc, e_cparhd' , e_minalloc, e_maxalloc, e_ss, e_sp, e_csum, e_ip, e_cs, e_lfarlc, e_ovno, e_oemid, e_oeminfo, e_lfanew
File Header	Machine, NumberOfSections, PointerToSymbolTable, NumberOfSymbols, SizeOfOptionalHeader, Characteristics
Optional Header	Magic, MajorLinkerVersion, MinorLinkerVersion, SizeOfCode, SizeOfInitializedData, SizeOfUninitializedData','AddressOfEntryPoint, BaseOfCode, ImageBase, SectionAlignment,FileAlignment, MajorOperatingSystemVersion, MinorOperatingSystemVersion, MajorImageVersion, MinorImageVersion, MajorSubsystemVersion, MinorSubsystemVersion, SizeOfHeaders, CheckSum, SizeOfImage, Subsystem, DllCharacteristics, SizeOfStackReserve, SizeOfStackCommit, SizeOfHeapReserve, SizeOfHeapCommit, LoaderFlags, NumberOfRvaAndSizes

B. *Methodology*

Untuk dapat mendeteksi malware menggunakan machine learning, diagram alur pada pembuatan model machine ditunjukan pada Gambar. (1) mengumpulkan dataset malware, (2) data preprocessing, (3) feature selection, dan (4) membuat model klasifikasi, dan analysis hasil berdasarkan metric.

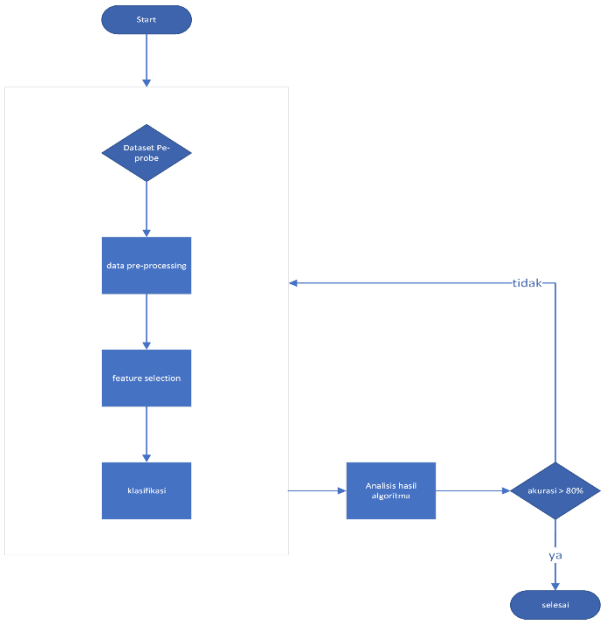

Figure 1: Alur metodologi pembuatan model machine learning untuk klasifikasi dan deteksi

IV. EXPERIMETNS AND RESULT

A. *Experiment*

Dalam penelitian ini untuk menentukan model machine learning terbaik dengan cara membagi dua subset dataset yang 80% untuk train dan 20% untuk testing dengan menggunakan kfold cross validation dengan nilai k = 10. Machine learning yang digunakan dalam klasifikasi dan deteksi malware yaitu Bagging dengan menggunakan parameter(tuning) dan Bagging tanpa menggunakan parameter(tuning).

Table 2 : Parameter(tuning) yang digunakan pada machine learning

Machine Learning	Parameter(tuning)
Bagging	Random_state = 0, N_estimator = 100, Max_Depth = 16, Max_features = sqrt

Confusion matrix digunakan untuk mengukur kinerja model klasifikasi machine learning. Tabel 3 menunjukan confusion matrix untuk klasifikasi malware, Di mana TP (True Positive) untuk mendeteksi file Malware, FN (False Negative) mendeteksi file bukan malware namun sebenarnya malware, FP (False Positive) mendeteksi file bukan malware namun dianggap malware, dan TN (True Negative) mendeteksi file bukan malware.

Table 3 : Confusion matrix untuk mengukur kinerja machine learning

		Actual	
		Positive	Negative
Predicted	Positive	True Positive (TP)	False Negative(FN)
	Negative	False Positive(FP)	True Negative(TN)

Dalam penelitian ini untuk mengukur kinerja machine learning berdasarkan confusion matrix digunakan rumus berdasarkan penelitian sebelumnya [14] dan [17] , berikut rumus yang digunakan untuk mengukur kinerja machine learning :

- Accuracy

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

- Detection Rate

$$DR = \frac{TP}{TP + FN}$$

- False Alarm Rate

$$FAR = \frac{FP}{FP + TN}$$

B. *Result Analysis*

Pada bagian ini, melakukan evaluasi model machine learning untuk deteksi malware, algoritma machine learning digunakan adalah Bagging dalam percobaan penelitian. Untuk mengukur kinerja machine learning digunakan rumus accuracy, detection rate, dan false alarm rate berdasarkan hasil confusion matrix setiap model machine learning.

Table 4 : Perfomansi klasifikasi machine learning tanpa menggunakan parameter(tuning)

Machine learning	Accuracy	Detection Rate	False Alarm Rate
Bagging tidak menggunakan parameter (tuning)	0.976	0.9881	0.062

Bagging menggunakan parameter(tuning)	0.976	0.9881	0.061

Berdasarkan hasil klasifikasi Tabel 4 machine learning Bagging dengan tidak menggunakan parameter(tuning) menghasilkan nilai accuracy sebesar 0.976, detection rate 0.9881, dan false alarm rate sebesar 0.062. Untuk machine learning Bagging yang menggunakan parameter menghasilkan nilai accuracy sebesar 0.976, detection rate 0.9881, dan false alarm rate sebesar 0.062.

Untuk menentukan penelitian yang dibuat sudah optimal dalam pembuatan model klasifikasi machine learning untuk deteksi malware, maka dibuat perbadingan dengan penelitian terdahulu, berikut perbadingan penelitian yang diusulkan dan penelitian yang sudah dibuat oleh para peneliti lain :

Table 5 : Perbandingan perfomansi machine learning terhdapat penelitian terdahulu dan penelitian yang dibuat

Methods	Machine Learning	Accuracy
Radwan, A. M., 2019	K-Nearest Neighbors (KNN), Gradient Boosted Trees (GPR), Decision Tree (DT), Random forest (RF), File large margin (FLM), Logistic Regression (LR) and Naïve Bayes (NB).	99.23%
Dong-Hee Kim et al., 2016	SVC dan Cart	99.99%
Vyas et al, 2017	KNN, Decision Tree, SVM, dan Random Forest	98.7%
Shafiq et al, 2009	J48	0.994%
Tina Razei et al., 2021	Random Forest	95.5%
Penelitian yang dibuat	Bagging	0.976%

Berdasarkan perbandingan pada Tabel 6 penelitian yang dilakukan Radwan, A, M [14] mengusulkan penelitian deteksi malware berdasarkan fitur dari dataset pe-file dengan menggunakan 6 model machine learning untuk melakukan deteksi dengan tingkat akurasi terbesar pada machine learning Random Forest dan NB. Selain itu Dong-Hee Kim[6] juga melakukan penelitian mengklasifikasikan file executable menjadi dua kategori, yaitu malware dan file benign dengan menggunakan algortima SVC dan CART, dengan menggunakan fitur static analysis.

Vyas [20] melakukan penelitian deteksi family malware dengan menggunakan metode berbasis fitur-fitur dari pe file, dengan mengklasifikasikan jenis malware yang berada didalam pe file. Selain itu Shafiq juga melakukan penelitian deteksi malware berdasarkan fitur dari pe file dengan cara membagi 2 jenis pe file antara packed dan non packed. Selain itu Tina Razei [15] juga membuat penelitian deteksi malware menggunakan 9 fitur hasil ekstrasi pe file dengan tinggkat akurasi tertinggi yaitu model machine learning.

Pada penelitian yang dibuat, penelitian ini berfokus kepada deteksi malware dengan menggunakan machine learning berdasarkan fitur-fitur ektrasi pe probe. Fitur yang digunakan pada penelitian ini yaitu bagian dari dos header, file header, dan optional header yang terdapat pada pe file dan mendapatkan akurasi tertinggi yaitu 0.976.

V. CONCLUSION AND FEATURE WORK

Pe probe merupakan format file yang paling banyak digunakan dalam system operasi windows, sehingga untuk mendeteksi malware yang menginfeksi pe probe sangat penting dilakukan. Untuk mengklasifikasikan malware menggunakan machine learning Bagging menggunakan parameter(tuning) dan tidak menggunakan parameter(tuning). Berdasarkan hasil pengujian machine learning Bagging mendapatkan nilai accuracy sebesar 0.976. Pada penelitian yang diusulkan untuk mentedeksi malware hanya terbatas pada file berformat pe probe(.exe dan .dll) dan machine learning yang dibuat hanya dapat mendeteksi malware berdasarkan feature dataset yang telah dipelajari. Untuk penelitian dimasa mendatang, diharapkan dapat membuat model deteksi malware yang dapat mendeteksi selain berformat pe probe.

ACKNOLEDGMENT

Penulis mengucapkan terima maksih kepada bapak Satria Mandala, PhD selaku dosen pembimbing yang telah membantu dalam penulisan penelitian ini.

REFERENCES

[1] Abijah Roseline, S., Geetha, S., Kadry, S., and Nam, Y. Intelligent Vision-based Malware Detection and Classification using Deep Random Forest Paradigm. IEEE Access (2020), 1–1.

[2] Chen, C. W., Su, C. H., Lee, K. W., and Bair, P. H. Malware Family Classification using Active Learning by Learning. International Conference on Advanced Communication Technology, ICACT 2020 (2020), 590–595.

[3] Chumachenko, K. Machine Learning Methods for Malware Detection and Classification. Proceedings of the 21st Pan-Hellenic Conference on Informatics - PCI 2017 (2017), 93.

[4] Das, S., Liu, Y., Zhang, W., and Chandramohan, M. Semanticsbased online malware detection: Towards efficient real-time protection against malware. IEEE Transactions on Information Forensics and Security 11, 2 (2016), 289–302.

[5] Han, X., Jin, F., Wang, R., Wang, S., and Yuan, Y. Classification of malware for self-driving systems. Neurocomputing428(2021),352–360.

[6] Kim, D., Woo, S., Lee, D., and Chung, T. Static detection of malware and benign executable using machine learning algorithm. In INTERNET 2016: The Eighth International Conference on Evolving Internet(2016),pp.14–19.

[7] Kshetri, N. Kaspersky lab: from russia with anti-virus. Emerald Emerging Markets Case Studies 1, 3 (2011),1–10.

[8] Liu, L., sheng Wang, B., Yu, B., and xi Zhong, Q. Automatic malware classification and new malware detection using machine learning. Frontiers of Information Technology and Electronic Engineering 18,9(2017),1336–1347.

[9] Liu, Y. S., Lai, Y. K., Wang, Z. H., and Yan, H. B. A New Learning Approach to Malware Classification Using Discriminative Feature Extraction. IEEE Access 7, c (2019), 13015–13023.

[10] Mangialardo, R. J., and Duarte, J. C. Integrating static and dynamic malware analysis using machine learning. IEEE Latin America Transactions 13, 9 (2015),3080–3087.

[11] Nataraj, L., Karthikeyan, S., Jacob, G., and Manjunath, B. S. Malware images: visualization and automatic classification. In Proceedings of the 8th international symposium on visualization for cyber security(2011),pp.1–7.

[12] Oh, S., Go, W., and Lee, T. A study on the behavior-based malware detection signature. In Advances on Broad-Band Wireless Computing, Communication and Applications: Proceedings of the 11th International Conference On Broad-Band Wireless Computing,Communication and Applications(BWCCA–2016) November 5–7, 2016, Korea (2017), Springer,pp.663–670.

[13] Pai, S., Troia, F. D., Visaggio, C. A., Austin, T. H., and Stamp,M. Clustering for malware classification. Journal of Computer Virology and Hacking Techniques 13, 2 (2017), 95–107.

[14] Radwan, A. M. Machine learning techniques to detect maliciousness of portable executable files. In 2019 International Conference on Promising Electronic Technologies (ICPET) (2019), IEEE, pp. 86–90.

[15] Rezaei, T., Manavi, F., and Hamzeh, A. A pe header-based method for malware detection using clustering and deep embedding techniques. Journal of Information Security and Applications 60 (2021), 102876.

[16] Sahu, M. K., Ahirwar, M., and Shukla, P. K. Improved malware detection technique using ensemble based classifier and graph theory. In 2015 IEEE International Conference on Computational Intelligence & Communication Technology (2015), IEEE, pp. 150-154.

[17] Shafiq, M. Z., Tabish, S., and Farooq, M. Pe-probe: leveraging packer detection and structural information to detect malicious portable executables. In Proceedings of the Virus Bulletin Conference (VB) (2009),vol. 8.

[18] Shafiq, M. Z., Tabish, S. M., Mirza, F., and Farooq, M. PE-Miner : Mining Structural Information to Detect Malicious Executables in Realtime Agenda Introduc1on to Domain Problem Defini1on Proposed Solu1on.121–141.

[19] Udayakumar, N., Saglani, V. J., Cupta, A. V., and Subbulakshmi, T. Malware Classification Using Machine Learning Algorithms. Proceedings of the 2nd International Conference on Trends in Electronics and Informatics, ICOEI 2018 (2018), 1007–1012

[20] Vyas, R., Luo, X., McFarland, N., and Justice, C. Investigation of malicious portable executable file detection on the network using super-vised learning techniques. In 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM) (2017), IEEE, pp. 941–946.

[21] Wuechner, T., Cislak, A., Ochoa, M., and Pretschner,A. Leveraging compression-based graph mining for behavior-based malware detection. IEEE Transactions on Dependable and Secure Computing 16,1 (2017), 99–112.

[22] Xia, S., Pan, Z., Chen, Z., Bai, W., and Yang, H. Malware classification with markov transition field encoded images. In 2018 Eighth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC) (2018), IEEE, pp. 1–5.

[23] Xue, D., Li, J., Lv, T., Wu, W., and Wang, J. Malware classification using probability scoring and machine learning. IEEE Access 7(2019), 91641–91656. 38

[24] Kumar, S., Janet, B., and Neelakantan, S. Identification of malware families using stacking of textural features and machine learning. Expert Systems with Applications 208 (2022), 118073.