BAB 6

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan hasil pengujian yang dilakukan pada Bab 5, maka didapat kesimpulan sebagai berikut:

1. *Web server Apache* memiliki beberapa direktori lagi di dalamnya seperti *cgi-bins, htdocs, conf, icons, images, sbin* dan *logs*. Direktori *logs* yang sering digunakan untuk membuka serta mengambil *file access log*.
2. *Log file* yang digunakan adalah *log* dengan format *Combined Log*. Format *Combined Log* tersebut terbagi menjadi alamat ip, *user-identifier*, *user id*, tanggal dan waktu, permintaan, kode status HTTP, ukuran objek, *referer*, dan *browser*.
3. *Webshell* yang digunakan untuk analisis ada lima yaitu, *webshell Webadmin, C99, WSO, R57*, dan *B374K*. Karakteristik dari masing-masing *webshell* adalah sebagai berikut:
   (a) *Webshell Webadmin* tercatat banyak baris *log* secara bersamaan untuk setiap aksi yang dilakukan. Terdapat beberapa parameter permintaan sehingga aksi yang dilakukan dapat terlihat. Parameter tersebut terdapat kata *image* dan *action*.
   (b) *Webshell C99* memiliki parameter permintaan yang berbeda namun mirip dengan *webshell Webadmin*, yaitu adanya kata *act, img*, dan *%2C*.
   (c) *Webshell WSO* dan *R57* tidak dapat diidentifikasi dan dibedakan dengan aktivitas normal pada *file access log*. *Webshell WSO* dan *R57* tidak memiliki karakter khusus sehingga mereka dijadikan satu kategori untuk mendeteksi aktivitas *webshell*.
   (d) *Webshell B374K* tidak begitu terlihat pada *access log* untuk aksi yang dilakukan, namun ada satu karakter yang bisa membedakannya dengan *webshell* lain, yaitu karakter tanda tanya (?), karakter |, dan karakter (-).
4. Pembangunan perangkat lunak untuk mendeteksi aktivitas *webshell* sudah berhasil dilakukan menggunakan model *machine learning*. Perangkat lunak yang dibangun menggunakan bahasa *Python* dan *framework Django* untuk membuat *web interface*. Perangkat lunak yang sudah dibangun dapat mendeteksi aktivitas *webshell* serta menampilkan laporan *log* untuk melihat *log* mana saja yang merupakan aktivitas normal dan aktivitas *webshell*.
5. Perangkat lunak berhasil mendeteksi *webshell Webadmin, C99,* dan *B374K* sedangkan gagal untuk mendeteksi *webshell R57* dan *WSO*. Hal ini dikarenakan karakteristik dari kedua *webshell* tersebut tidak dapat diidentifikasi dan dibedakan.
6. Metode pembelajaran mesin yang digunakan pada perangkat lunak untuk mendeteksi aktivitas *webshell* adalah metode *Decision Tree*.
7. *Log* yang diunggah ke dalam *web* merupakan *data train* yang diolah kemudian dibentuk model *Decision Tree*. *Data train* yang sudah diolah tersebut menjadi acuan untuk mendeteksi aktivitas *webshell* pada *log* yang diuji dan dieksperimen. Perbandingan hasil prediksi dan hasil preproses pada *log* yang diuji menghasilkan angka akurasi. Angka akurasi merepresentasikan seberapa akurat prediksi aktivitas *webshell* dengan model *Decision Tree*.

6.2 Saran

Berdasarkan kesimpulan yang dibuat, terdapat beberapa saran yang dapat digunakan di penelitian selanjutnya:

1. Dapat menganalisis *webshell* lainnya selain *webshell Webadmin, C99, WSO, R57,* dan *B374K.* Hal ini dikarenakan masih banyak jenis *webshell* lain yang dapat dicoba.
2. Menggunakan metode *machine learning* lain untuk mendeteksi adanya aktivitas *webshell.*
3. *Web interface* yang dibuat dapat lebih interaktif, memiliki tampilan yang lebih bagus, dan memiliki banyak fitur.
4. Mencoba *log* untuk diuji dari berbagai macam *web server* selain *apache*, seperti *Nginx* dan *Apache Tomcat.*

DAFTAR REFERENSI

[1] Mavridis, I. dan Karatza, E. (2015) Log file analysis in cloud with apache hadoop and apache spark. *Proceedings of the Second International Workshop on Sustainable Ultrascale Computing Systems (NESUS 2015)*, Krakow, Poland, October, pp. 51–62. Universidad Carlos III de Madrid.

[2] Li, Y., Jiang, Y., Gu, J., Lu, M., Yu, M., Armstrong, E. M., Huang, T., Moroni, D., McGibbney, L. J., Frank, G., dan Yang, C. (2019) A cloud-based framework for large-scale log mining through apache spark and elasticsearch. *Applied Sciences*, **9**, 1–13.

[3] Guo, Y., Marco-Gisbert, H., dan Keir, P. (2020) Mitigating webshell attacks through machine learning techniques. *Future Internet*, **12**, 12.

[4] McKinney, W. (2017) *Python for Data Analysis*, 2nd edition. OŔeilly Media, Inc, Sebastopol, California.

[5] Kurose, J. F. dan Ross, K. W. (2013) *Computer Networking : A Top Down Approach 6th Edition*, 6th edition. Library of Congress Cataloging-in-Publication Data, United States of America.

[6] Cahyanto, T. A. dan Prayudi, Y. (2014) Investigasi forensika pada log web server untuk menemukan bukti digital terkait dengan serangan menggunakan metode hidden markov models. *SNATI 2014*, Yogyakarta, 21 Juni, pp. 15–19. Seminar Nasional Aplikasi Teknologi Informasi (SNATI).

[7] Koprawi, M. (2020) Dampak dan pencegahan serangan file inclusion: Perspektif developer. *Jurnal Nasional Informatika dan Teknologi Jaringan*, **4**, 91–95.

[8] Kotsiantis, S. (2007) Supervised machine learning: A review of classification. *Emerging Artificial Intelligence Applicants in Computer Engineering*, **160**, 3–24.

[9] Liu, D. (2009) *Cisco Router and Switch Forensics*, 1st edition. Syngress, Boston.

[10] Chandra, A. (2019) Analisis performansi antara apache & nginx web server dalam menangani client request. *Journal of Systems Integration*, **14**, 48–56.

[11] Grace, L. J., V.Maheswari, dan Nagamalai, D. (2011) Analysis of web logs and web user in web mining. *International Journal of Network Security & Its Applications (IJNSA)*, **3**, 99–110.

[12] Hassan, M., Bhuyian, T., Sohel, M. K., Sharif, H., dan Biswas, S. (2018) Saisan: An automated local file inclusion vulnerability detection model. *International journal of engineering and technology*, **7**, 4.

[13] Wu, Y., Sun, Y., Huang, C., Jia, P., dan Liu, L. (2019) Sesion-based webshell detection using machine learning in web logs. *Hindawi, Security and Communication Networks*, **2019**.

[14] Eri, H. (2017) Analisis forensik malicious software wso webshell pada platform linux. *Universitas Janabadra*, **1**, 1–11.

[15] Agisilaos, S. (2016) Detecting malicious code in a web server. Thesis. UNIVERSITY OF PIRAEUS, Athena.

[16] Mitchell, T. M. (1997) *Machine Learning*, 1st edition. McGraw-Hill Science, United States of America.

[17] Breiman, L., Friedman, J., Olshen, R., dan Stone, C. J. (1983) Classification and regression trees. *Biometrics*, September 874. International Biometric Society.

[18] Hamoud, A. K., Hashim, A. S., dan Awadh, W. A. (2018) Predicting student performance in higher education institutions using decision tree analysis. *Int. J. Interact. Multim. Artif. Intell.*, **5**, 26–31.

[19] Nasution, D. A., Khotimah, H. H., dan Chamidah, N. (2019) Perbandingan normalisasi data untuk klasifikasi wine menggunakan algoritma k-nn. *Computer Engineering and Science*, **4**, 78–82.

[20] Hasnain, M., Pasha, M. F., Ghani, I., Imran, M., Alzahrani, M. Y., dan Budiarto, R. (2020) Evaluating trust prediction and confusion matrix measures for web services ranking. *IEEE Access*, **8**, 90847–90861.