International Journal of Advanced Technology in Engineering and Science
Vol. No.08, Issue No. 09, September 2020
www.ijates.com
ijates
ISSN 2348 - 7550

Advanced Identification Method for Network Anomaly Using Data Mining Techniques

Bolla. Ramesh Babu M.Tech **K. Gopi** M.Tech **I. Sriram Murthy** M.Tech

Assistant ProfessorAssistant ProfessorAssistant Professor

ABSTRACT

There is now a huge and growing concern among the scientific community regarding information and communication technology (ICT) security because any attack or network anomaly can have a significant impact on many domains, such as national security, private data storage, social welfare, economic problems, and so on. The domain of detection of anomalies is therefore a wide area of study, and several different techniques and methods have developed over the years to this end. Attacks, problems, and internal failures can damage a whole network system if not detected early.Thus this study introduces an autonomous profile-based anomaly detection system based on the Principal Component Analysis (PCADS-AD) statistical method. This approach generates a network profile called Digital Signature of Network Segment using Flow Analysis (DSNSF) that, through historical data analysis, denotes the expected normal behavior of a network traffic operation. The digital signature is used to detect anomalies in the normal traffic trend as a criterion for volume anomaly detection.Seven traffic flow characteristics are used for the proposed system: bits, packets, and flow numbers to identify problems, and IP addresses and ports from source to destination to provide the network administrator with the necessary details to fix them. The observed findings seek to contribute to the development of state-of-the-art approaches and anomaly detection strategies that aim to address certain challenges emerging from the continuous growth in complexity, speed, and size of today's large-scale networks, as well as producing high-value findings for better real-time detection.

International Journal of Advanced Technology in Engineering and Science
Vol. No.08, Issue No. 09, September 2020
www.ijates.com
ijates
ISSN 2348 - 7550

Keywords: Anomaly Detection, Intrusion Detection System, Network Security, Principal Component Analysis

1. INTRODUCTION

The research community is now continuously concerned about stability and service quality in high-efficiency networks. Growing the number of connected networking computers, site users, utilities, and apps, the expansion of modern networking technology and software tends to make computer networks bigger and more flexible as structures. Moreover, for next-generation networks, there is a so-called unbounded connectivity model, which envisages delivering connectivity to users anywhere, anytime, and in full integration and interoperability of emerging technologies [1]. These challenges make ensuring correct network maintenance which leads to significant network flaws much more complicated which complicated as safety incidences can occur more regularly. [2, 3].

Such protection instances may result from malicious attacks aimed at locking down networks or stealing private information or from internal (operational) causes, such as configuration failures, server crashes, power outages, network congestion, or large-file non-malicious transfers [4]. Whatever the cause, attacks, generally referred to as irregularities, may have a huge effect on network service and end-users, as well as the activity and availability of computer networks.A multi-definition word paradox. Barnett and Lewis describe anomaly of the data collection as "observation (or a subset of observations) which seems incompatible with the rest of the data collection"[5]. This concept was characterized by Chandola et al. as "data patterns in which a notion of normal behavior has been not clearly defined[6]." "Anomalies are rare and substantial variations in the amount of traffic of a network that can also cover several connexions," according to Lakhina and others[7]. Hoque et al. describe it as "significant phenomena that are not in line with the well-defined definition of the norm"[8].

2. RELATED WORK

In the science community, there are already tremendous and increasing worries regarding ICT security, since attacks and anomaly on the network may have a major impact on many fields, including national security, privately-owned storage of data, social services, economic problems, and others. The field of identification of anomalies is thus a wide field of study, and a great many diverse methods and methods have developed over the years.Since the beginning of the 19th century, researchers have been researching the topic of anomaly detection and have generated a significant number of papers using different methods, from mathematical simulation to evolutionary computing approaches. However, defining and categorizing all anomaly detection methods is not a simple feat.A broad variety of concepts, such as event types, device types, approaches and algorithms used, and technological

challenges such as transmission costs and network performance, should be addressed. That is why many works attempt to outline many of these items but are incapable of presenting the wider image of the range for anomaly detection.

As in[18] and[9], the emphasis is just on the most common techniques and approaches like machine learning, categorization, and statistics. She also addressed the entire issue statement briefly and described core issues such as data sets, problems, and guidelines, such as[19] and [20]. The anomaly analysis of backbone networks was checked by Marnerides et al [21]. Although several main topics related to anomaly detection are outlined in all those surveys, they are not completely complete.Some stress anomaly types but do not address all sorts of techniques, for example, whereas some are studying common approaches, while they overlook the nature of intrusion detection systems and data entry, etc. The description of context analysis, as well as a core review of the applicable procedures, processes and structures in the field, will also be investigated as the key aim for anomaly detection.To make it simpler to understand the framework of this chapter, five dimensions were taken of the domain of the detecting anomalies: (i) traffic anomalies in the network, (ii) network data types, (iii) categories of intrusion detection systems, (iv) methods and systems of detection and (v) accessible problem.

A significant factor in the identification strategy is the existence of an anomaly. It may or may not be an abnormality, depending on the sense in which an abnormality happens or how it happens. This is how the device treats and recognizes irregularities that have been observed and identified. There are three types of anomalies, based on their nature: anomalies of the points, cumulative anomalies, and history anomalies [6, 10, 22].

3. PROPOSED SYSTEM

In this chapter, the hybrid anomaly detection system using principal component analysisis presented. However, before explaining its full process, Figure 1 summarizes the overalloperation of it. The system is divided into two parts: Traffic Characterization and Anomaly Detection.The traffic characterization is responsible for extracting quantitative attributes (bits/s,packets/s, and several flows/s) from a flow database containing historical data about thenetwork segment activity, and generate the corresponding DSNSFs.The mentioned components are deployed in conjunction with one another to filter packets on the communication networks, such as mobile networks, and for certain network protocols that are known or considered to be vulnerable to or used in cyber-attacks. This allows the HADM to expend a smaller amount of processing resources on other network protocols, such as streaming protocols that are not normally vulnerable and thus not typically targeted by cyber-attackers. The ability of the HADM to focus on vulnerable network protocols helps to avoid burdening network servers with the unnecessary computational load. The protocol analyzer filters the network packets

International Journal of Advanced Technology in Engineering and Science
Vol. No.08, Issue No. 09, September 2020
www.ijates.com
ijates
ISSN 2348 - 7550

and identifies vulnerable protocols. The non-vulnerable protocols are forwarded to the feature extraction module for further processing. The feature extraction module extracts features from the incoming packets and provides these features to the learning algorithm I for the analysis. If the output from the learning algorithm is suspicious, it is recorded into a log file. If traffic is carried on the vulnerable protocol, the counter and prioritization module forwards the suspicious traffic to the next level based on the occurrence of the protocol against a defined threshold.

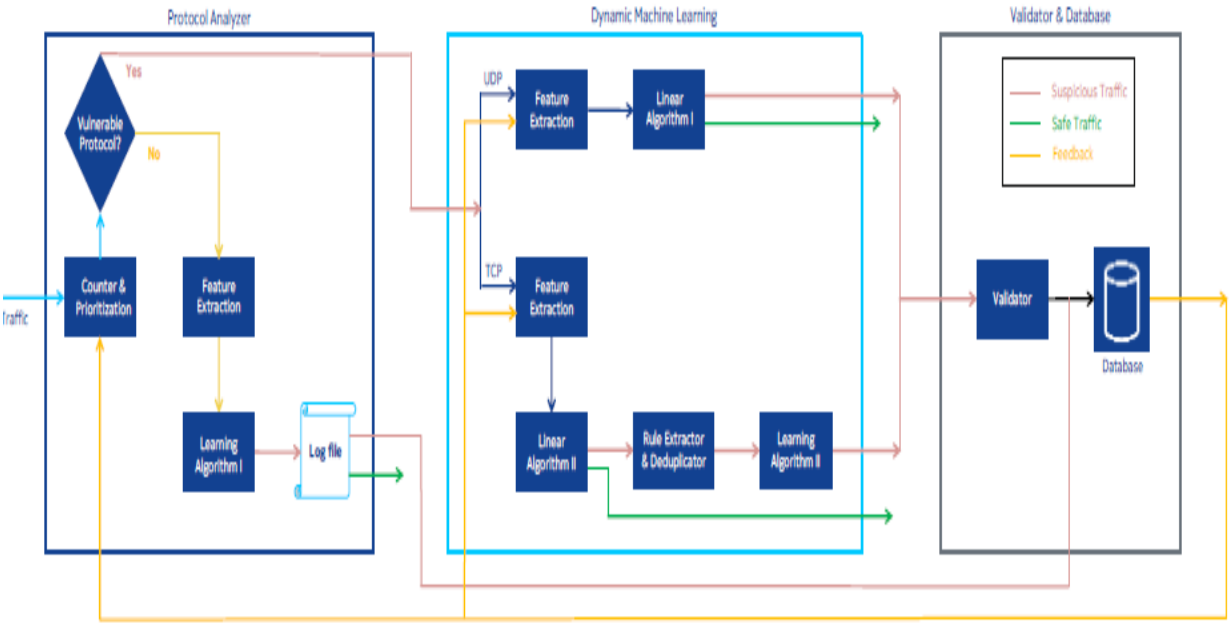

Figure 1. Proposed system architecture

4. RESULTS

The new system was developed from the notion of Hybrid PSO and C4.5. The structure of IDS is expressed in the ideas of the SKNN Classifier modified in R in this study. The "klaR" bundle available in R is in this job. The acquired results show ample accuracy. The findings have been shown.

International Journal of Advanced Technology in Engineering and Science
Vol. No.08, Issue No. 09, September 2020
www.ijates.com
ijates
ISSN 2348 - 7550

Table 1: Results Comparison

Techniques	Sensitivity	Specificity	Accuracy	FAR
C4.5	87.57	83	91.24	1.45
SVM	81.92	63.29	88.27	3.01
C4.5+ACO	89.15	86.43	96.15	0.88
SVM+ACO	97.31	69.66	91.82	1.11
C4.5+PSO	93.40	89.88	96.37	1.83
SVM+PSO	91.50	71.10	92.59	2.96
EDADT	96.65	92.25	97.11	0.20
Proposed Method	99.81	99.90	99.62	0.01

Using the SKNN Model classifying and statistical analytics method the proposed model has been developed, the R programming language is used for analytical and classification activities. The library kit of KLAR will conform to the classification of various labels used. Figure 2 presents the findings of the identification of irregularities and mismanagement.

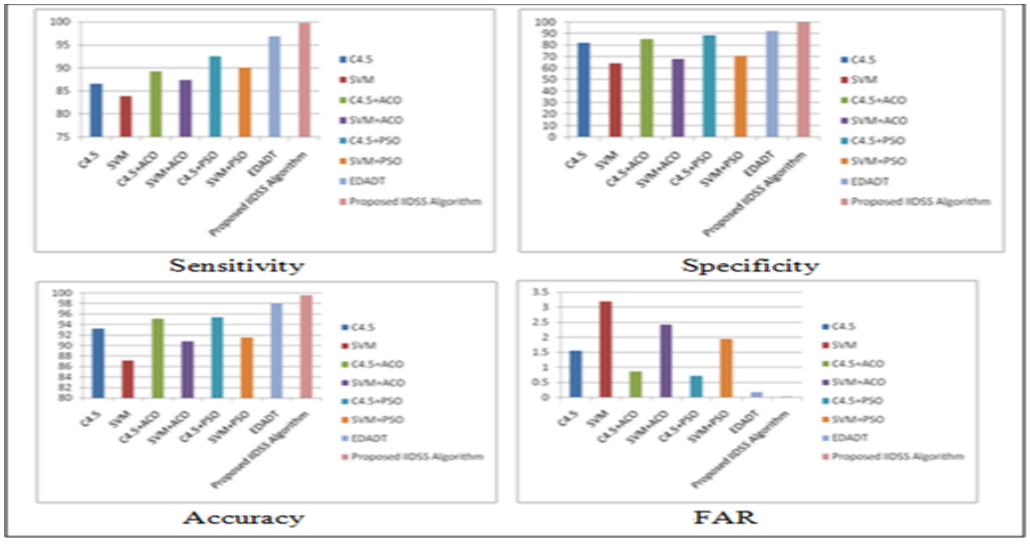

Figure 2: Results observed -Sensitivity, Specificity,and Accuracy and FAR

CONCLUSION

Better performance than current techniques were suggested by the proposed intrusion detection monitor. The technique has demonstrated wide efficiency by restricting the faulty warning rate and decreases the remaining pressure on managers. In comparison, this model has increased its effect by 13.24% and C4.5, by 10.55% in comparison and C4.5+ACO, while

comparing and EDADT have increased by 2.85%. The effects of research display greater comparing the accuracy and the established system. The Future Study is planned to modify the IDS to differentiate the number of attacks and to extend the tally from 23 to 40.This study showed that warnings are produced if the behavioral pattern of the packets differs. The patterns are aligned to the suggested basis of signature rules for snort. In comparison to the current snort rules, the new scheme has been reviewed methodically and the proposed rules have proven more detailed and reliable. Advanced data processing tools and learning machines used to spot new malicious attacks on a vast volume of data would be used in future work.

REFERENCES

1. A.Saidi et al.:, The functional of A Mobile Agent System to Enhance DoS and DDoS Detection in Cloud, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 6 (2016) pp 4615-4617
2. Adeeb Alhomoud et al., Performance Evaluation Study of Intrusion Detection Systems, The 2nd International Conference on Ambient Systems, Networks and Technologies, (ANT), Procedia Computer Science 5 (2011) 173–180, 1877–0509 © 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of Prof. Elhadi Shakshuki and Prof. Muhammad Younas. DOI:10.1016/j.procs.2011.07.024
3. Anderson, James P., "Computer Security Threat Monitoring and Surveillance," Washing, PA, James P. Anderson Co., 1980.
4. Anna L. Buczak. (2015). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection, 10.1109/COMST.2015.2494502, IEEE Communications Surveys & Tutorials, 1553-877X (c)
5. Aymen Abid et al.:, Outlier detection for wireless sensor networks using a density-based clustering approach, IET Wireless. Sens. Syst., 2017, Vol. 7 Iss. 4, pp. 83-90, The Institution of Engineering and Technology 2017, ISSN 2043-6386
6. Bellovin, S.M. "Network Firewalls", IEEE Communications Magazine, Vol. 32, pp. 50- 57, 1994.
7. Berchtold, B. Ertl, D. A. Keim, H.-P. Kriegel, and T. Seidl. Fast nearest neighbor search in high-dimensional space. In Proceedings of the Fourteenth International Conference on Data Engineering, ICDE '98, pages 209–218, Washington, DC, USA, 1998. IEEE Computer Society.
8. Blum, Avrim L. & Pat Langley (1997). Selection of relevant features and examples in machine learning. Artificial Intelligence, 97(1-2), 245–271
9. Catania Carlos A, Garino Carlos. Automatic network intrusion detection: current techniques and open issues. Elsevier Comput Electr Eng 2012; 38(5):1062–72.

10. Chien-Yi Chiu, Yuh-Jye Lee, Chien-Chung Chang. Semi-supervised learning for false alarm reduction. In: Industrial conference on data mining, no. 10; 2010. p. 595–605.
11. Ching-Hao, Hahn-Ming L, Devi P, Tsuhan C, Si-Yu H. Semi-supervised co-training, and active learning-based approach for multi-view intrusion detection. In: ACM symposium on applied computing, no. 9; 2009. p. 2042– 7.
12. Claude Turner et al. (2016). A Rule Status Monitoring Algorithm for Rule-Based Intrusion Detection and Prevention Systems, Complex Adaptive Systems, Conference Organized by Missouri University of Science and Technology 2016 - Los Angeles, CA, Procedia Computer Science 95 (2016) 361 – 368, 1877-0509, DOI: 10.1016/j.procs.2016.09.346
13. Das, S. (2001). Filters, Wrappers, and a Boosting-Based Hybrid for Feature Selection. Proc. 18thInt'l Conf. Machine Learning, 74-81
14. Dasgupta, D., and F. A. Gonzalez, "An intelligent decision support system for intrusion detection and response", In Proc. Of International Workshop on Mathematical Methods, Models, and Architectures for Computer Networks Security (MMM-ACNS), St.Petersburg. Springer- , 21-23 May,2001.
15. Denning, D.E. "An Intrusion-Detection Model", in IEEE Transactions on Software Engineering, Vol.13, No. 2, pp. 222-232, 1987.
16. Dickerson, J. E., and J. A. Dickerson, "Fuzzy network profiling for intrusion detection", In Proc. of NAFIPS 19th International Conference of the North American Fuzzy Information Processing Society, Atlanta, pp. 301306. North American Fuzzy Information
17. Divya and Surendra Lakra, "SNORT: A Hybrid intrusion detection system using artificial intelligence with a snort", International journal computer technology & application, Vol 4(3), 466-470, 2013.
18. E.Kesavulu Reddy, Member IAENG, V.Naveen Reddy, P.Govinda Rajulu. A Study of Intrusion Detection in Data Mining. Proceedings of the World Congress on Engineering 2011 Vol III WCE 2011, July 6 - 8, 2011, London, U.K.
19. Eskin, E., Arnold, A., Prerau, M., Portnoy, L., and Stolfo, S. J., A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data, In D.Barbarà and S. Jajodia (eds.), Applications of Data Mining in Computer Security, Kluwer Academic Publishers, Boston, MA, 2002, pp. 78-99.
20. Fayyad, U. M., G. Piatetsky-Shapiro, and P. Smyth, "The KDD process for extracting useful knowledge from volumes of data," Communications of the ACM 39 (11), November 1996, 2734.