International Journal on Applications in Engineering and Technology
Volume 8 : Issue 2 : May 2022, pp 87 – 92 www.aetsjournal.com
ISSN (Online) : 2455 - 0523

INTRUSION DETECTION SYSTEM

SURYARAJ R[1], SURENDAR U[2], SIVAKUMAR S[3], R MURUGESAN [4]

[1, 2, 3] Undergraduate student, Department of Computer Science and Engineering, Pavai College of Technology
[4] Assistant Professor, Department of Computer Science and Engineering, Pavai College of Technology

Abstract: *-* **–** Intrusion Detection System (IDS) defined as a Device or software application which monitors the network or system activities and finds if there is any malicious activity occur. Outstanding growth and usage of internet raises concerns about how to communicate and protect the digital information safely. In today's world hackers use different types of attacks for getting the valuable information. Many of the intrusion detection techniques, methods and algorithms help to detect those several attacks. The main objective of this paper is to provide a complete study about the intrusion detection, types of intrusion detection methods, types of attacks, different tools and techniques, research needs, challenges and finally develop the IDS Tool for Research Purpose That tool are capable of detect and prevent the intrusion from the intruder. This is for use in anomaly-detection based intrusion detection system. Containers provide isolation between the host system and the containerized environment by efficiently packaging applications along with their dependencies. This way, containers become a portable software environment for applications to run and scale. Unlike virtual machines, containers share the same kernel as the host operating system. This is leveraged to monitor the system calls of the container from the host system for anomaly detection. Thus, the monitoring system is not required to have any knowledge about the container nature, neither does the host system or the container being monitored need to be modified..

Key words: IDS, anomaly detection, network threat, container environment, Docker hub, Redis server, slips.

I. INTRODUCTION

Intrusion is the unauthorized access to a system's data or resources by an unauthorized user, the masquerader, or a software program. It is a threat faced by computer systems every day. Intrusion detection systems (IDS) and Intrusion Prevention Systems (IPS) are used to detect and prevent intrusions. IDS monitor networks and host systems to detect anomalies in system behavior, and abuse of system resources or policies. This includes collecting system usage information, and creating profiles based on normal and anomalous usage patterns . Profiles can be created based on a variety of parameters like CPU usage, login time, application usage order, login location, login time, session times (start and end), commands issued, system calls etc. If an abuse or anomalous activity that does not match the normal usage profiles is detected, then the session can be reported as an intrusion.

1.1 Machine learning in IDS

IDS is a software application to detect network intrusion using various machine learning algorithms.IDS monitors a network or system for malicious activity and protects a computer network from unauthorized access from users, including perhaps insider. The intrusion detector learning task is to build a predictive model (i.e. a classifier) capable of distinguishing between 'bad connections' (intrusion/attacks) and a 'good (normal) connections'.

1.2 IDS Architecture

Intrusion Detection System (IDS) is renowned and widely-deployed security tool to detect attacks and malicious activities in information system. It is generally deployed as a second line of defense along with other defensive security mechanisms, such as firewall, vulnerability monitor, access control and authentication that protects information system. Intrusion detection involves monitoring network traffic, detecting attempts to gain unauthorized access to a system or resources, and alerting the appropriate persons so that countermeasures can be taken. It also involves developing an understanding of how attacks occur. IDS is an essential element of any modern information system.

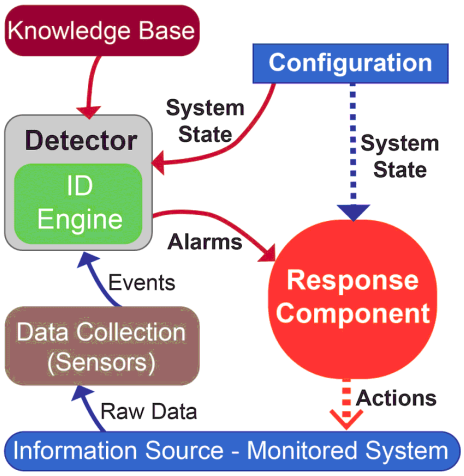


Fig.1.Overview of IDS architecture

II. DOCKER

Containers, or Linux Containers, are a technology that allows us to isolate certain kernel processes and trick them into thinking they're the only ones running in a completely new computer.Different from Virtual Machinesa container can share the kernel of the operating system while only having their different binaries/libraries loaded with them.In other words, you don't need to have whole different OS (called **guest OS**) installed inside your host OS. You can have several containers running within a single OS without having several different guest OS's installed.Docker packages an application and all its dependencies in a virtual container that can run on any Linux server. This is why we call them containers. Because they have all the necessary dependencies contained in a single piece of software.

2.1 The CTU-13 Dataset. A Labeled Dataset with Botnet, Normal and Background traffic

The CTU-13 is a dataset of botnet traffic that was captured in the CTU University, Czech Republic, in 2011. The goal of the dataset was to have a large capture of real **botnet** traffic mixed with **normal** traffic and **background** traffic. The CTU-13 dataset consists in thirteen captures (called **scenarios**) of different botnet samples. On each scenario we executed a specific malware, which used several protocols and performed different actions. Table shows the characteristics of the botnet scenarios.

Fig.1.1Characteristics of botnet

Table 2 – Characteristics of the botnet scenarios. (CF: ClickFraud, PS: Port Scan, FF: FastFlux, US: Compiled and controlled by us.)

Id	IRC	SPAM	CF	PS	DDoS	FF	P2P	US	HTTP	Note
1	√	√	√							
2	√	√	√							
3	√			√				√		
4	√				√			√		UDP and ICMP DDoS.
5		√		√					√	Scan web proxies.
6				√						Proprietary C&C. RDP.
7									√	Chinese hosts.
8				√						Proprietary C&C. Net-BIOS, STUN.
9	√	√	√	√						
10	√				√			√		UDP DDoS.
11	√				√			√		ICMP DDoS.
12							√			Synchronization.
13		√		√					√	Captcha. Web mail.

Each scenario was captured in a pcap file that contains all the packets of the three types of traffic. These pcap files were processed to obtain other type of information, such as NetFlows, WebLogs, etc. The first analysis of the CTU-13 dataset, that was described and published in the paper "An empirical comparison of botnet detection methods" (see Citation below) used unidirectional NetFlows to represent the traffic and to assign the labels. These unidirectional NetFlows should not be used because they were **outperformed** by our second analysis of the dataset, which used bidirectional NetFlows. The bidirectional NetFlows have several advantages over the directional ones. First, they solve the issue of differentiating between the client and the server, second they include more information and third they include **much more detailed-labels.**

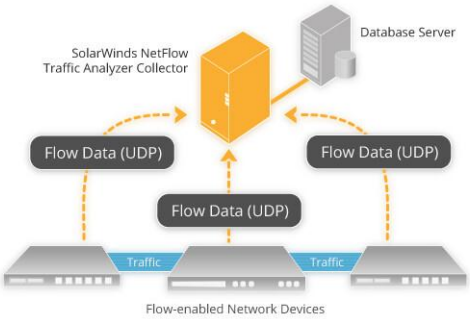


Fig .2 Statistics of Netflows

Scen.	Total Flows	Botnet Flows	Normal Flows	C&C Flows	Background Flows
1	2,824,636	39,933(1.41%)	30,387(1.07%)	1,026(0.03%)	2,753,290(97.47%)
2	1,808,122	18,839(1.04%)	9,120(0.5%)	2,102(0.11%)	1,778,061(98.33%)
3	4,710,638	26,759(0.56%)	116,887(2.48%)	63(0.001%)	4,566,929(96.94%)
4	1,121,076	1,719(0.15%)	25,268(2.25%)	49(0.004%)	1,094,040(97.58%)
5	129,832	695(0.53%)	4,679(3.6%)	206(1.15%)	124,252(95.7%)
6	558,919	4,431(0.79%)	7,494(1.34%)	199(0.03%)	546,795(97.83%)
7	114,077	37(0.03%)	1,677(1.47%)	26(0.02%)	112,337(98.47%)
8	2,954,230	5,052(0.17%)	72,822(2.46%)	1,074(2.4%)	2,875,282(97.32%)
9	2,753,884	179,880(6.5%)	43,340(1.57%)	5,099(0.18%)	2,525,565(91.7%)
10	1,309,791	106,315(8.11%)	15,847(1.2%)	37(0.002%)	1,187,592(90.67%)
11	107,251	8,161(7.6%)	2,718(2.53%)	3(0.002%)	96,369(89.85%)
12	325,471	2,143(0.65%)	7,628(2.34%)	25(0.007%)	315,675(96.99%)
13	1,925,149	38,791(2.01%)	31,939(1.65%)	1,202(0.06%)	1,853,217(96.26%)

III. Slips v0.9.1

Slips is a Python-based intrusion prevention system that uses machine learning to detect malicious behaviors in the network traffic. Slips was designed to focus on targeted attacks, to detect of command and control channelsi, and to provide good visualisation for the analyst. Slips is able to analyze real live traffic from the device and the large network captures in the type of a pcap files, Suricata, Zeek/Bro and Argus flows. As a result, Slips highlights suspicious behaviour and connections that needs to be deeper analyzed.

.

3.1 P2P

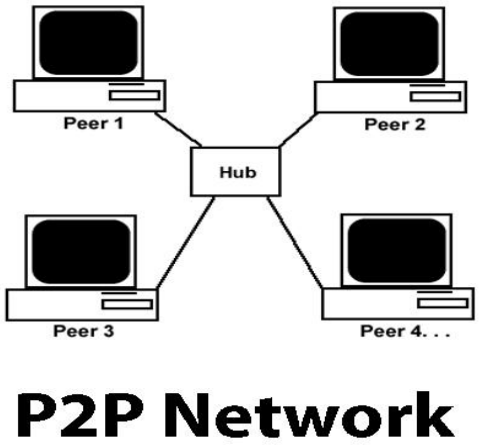


Fig -3: Name of the figure

The P2P module makes Slips be a peer in a peer to peer network of computers in the local network. The peers are only in the local network and they communicate using multicast packets. The P2P module is a highly complex system of data sharing, reports on malicious computers, asking about IoC to the peers and a complex trust model that is designed to resiste adversarial peers in the network. Adversarial peers are malicious peers that lie about the data being shared (like saying that a computer is maliciuos when is not, or that an attacker is benign).

Features of P2P network:

These networks do not involve a large number of nodes, usually less than 12. All the computers in the network store their own data but this data is accessible by the group. Unlike client-server networks, P2P uses resources and also provides them. This results in additional resources if the number of nodes increases.It requires specialized software. It allows resource sharing among the network. Since the nodes act as servers also, there is a constant threat of attack. Almost all the OS today support P2P networks.

3.2 Random Forest Algorithm

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. As the name suggests, **"Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset."** Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.**The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.**

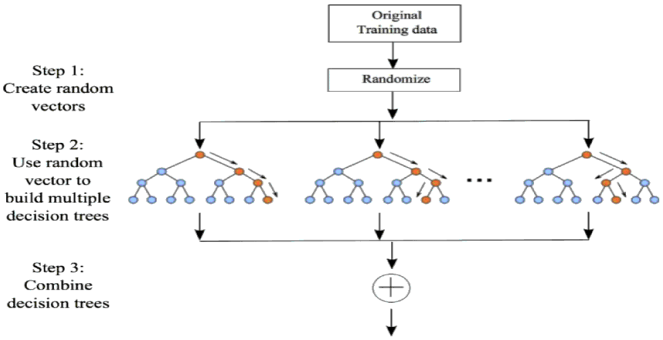


Fig.1.4 working of RF algorithm

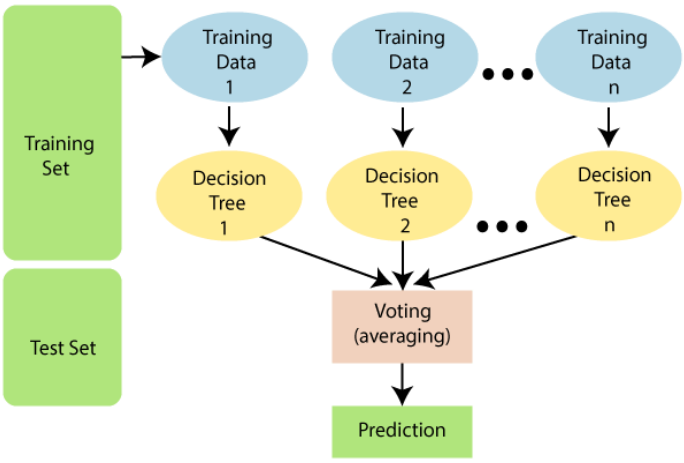


Fig.1.5 Training datasets of decision tree

IV. ZEEK NETWORK ANALYZER

Zeek is **a passive, open-source network traffic analyzer**. Many operators use Zeek as a network security monitor (NSM) to support investigations of suspicious or malicious activity. Zeek also supports a wide range of traffic analysis tasks beyond the security domain, including performance measurement and troubleshooting.

4.1 ZEEK FILE LOGS

One of Zeek's powerful features is the ability to extract content from network traffic and write it to disk as a file. This is easiest to understand with a protocol like File Transfer Protocol (FTP), a classic means to exchange files over a channel separate from that used to exchange commands. Protocols like HTTP are slightly more

complicated, as it includes headers which must be interpreted and not included in any file content transferred by the protocol.

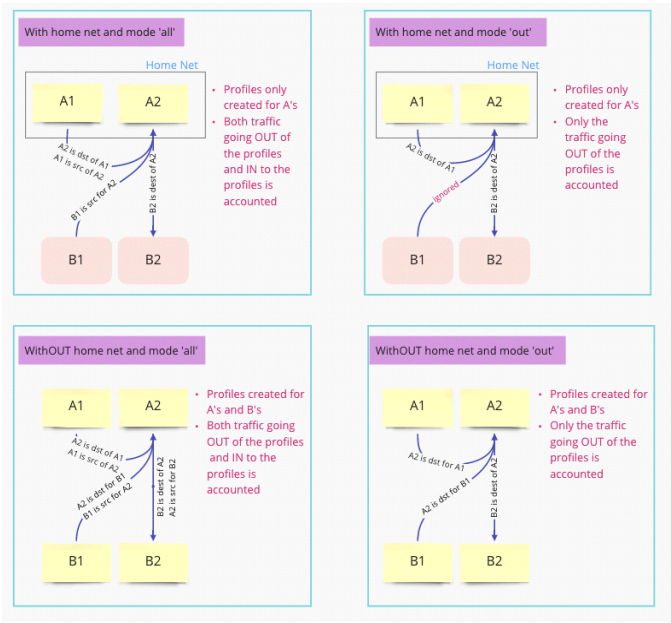
Zeek's files.log is a record of files that Zeek observed while inspecting network traffic. The existence of an entry in files.log does not mean that Zeek necessarily extracted file content and wrote it to disk. Analysts must configure Zeek to extract files by type in order to have them written to disk.an analyst has configured Zeek to extract files of MIME type application/x-dosexec and write them to disk. To understand the chain of events that result in having a file on disk, we will start with the conn.log, progress to the http.log, and conclude with the files.log.

The Zeek scripting manual, derived from the Zeek source code, completely explains the meaning of each field in the files.log (and other logs). It would be duplicative to manually recreate that information in another format here.

Therefore, this entry seeks to show how an analyst would make use of the information in the files.log

4.2 ANALYSIS DIRECTION

analysis_direction can either be out or all



4.3 YARA TOOL

YARA is the name of **a tool primarily used in malware research and detection**. It provides a rule-based approach to create descriptions of malware families based on textual or binary patterns. A description is essentially a YARA rule name, where these rules consist of sets of strings and a boolean expression

Writing YARA rules

YARA rules are easy to write and understand, and they have a syntax that resembles the C language. He here is the simplest rule that you can write for YARA, which does absolutely nothing:

```
rule dummy
{
   condition:
      false
}
```

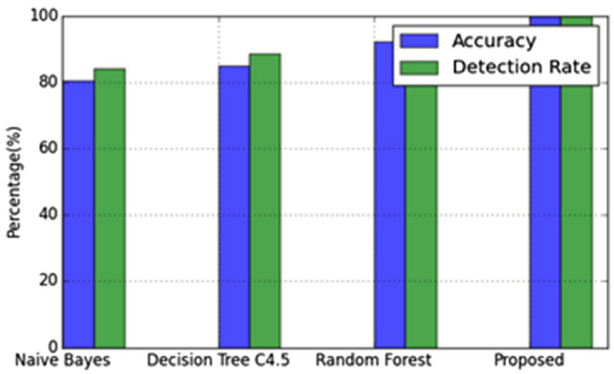
Each rule in YARA starts with the keyword rule followed by a rule identifier. Identifiers must follow the same lexical conventions of the C programming language, they can contain any alphanumeric character and the underscore character, but the first character can not be a digit. Rule identifiers are case sensitive and cannot exceed 128 characters.

4.4 REDIS SERVER

Red1is is an open source, advanced key-value store and an apt solution for building high performance, scalable web applications.
Redis has three main peculiarities that sets it apart.

- Redis holds its database entirely in the memory, using the disk only for persistence.
- Redis has a relatively rich set of data types when compared to many key-value data stores.
- Redis can replicate data to any number of slaves.

4.5 ACCURACY AND DETECTION OF RANDOM FOREST



V. CONCLUSION

Intrusion Detection System plays a very important and vital role in the field of network security. The performance of the classifier is degrading by using intrusive patterns, accuracy and also, it's time-consuming. Random Forest, Neural Network are the ML and DL algorithms considered for IDS. The CTU-13 dataset used to value these accounts. With the CTU-13 dataset's aid, we proposed a DL and ML approach. By looking at the results that depend on accuracy, the powerful classifier can be identified. It seems reasonable that the random forest classifier outperforms better with accuracy of 98.98% based on the results.

An IDS uses very good signature analysis mechanisms to detect intrusions or potential misuse; however, organizations must still ensure that they have strong user identification and authentication mechanism in place. An IDS is not a substitute for a good security policy: As with good security and monitoring products, an IDS functions is one element of a corporate security policy. Successful intrusion detection requires that a well defined policy must be followed to ensure that vulnerabilities, intrusions and virus outbreaks, etc. are handled according to corporate security policy guidelines. Human intervention is required: The security administrator or network manager must investigate the attack once. It is detected and reported, determine how it has occurred, correct the problem and take the necessary actions to prevent the occurrences of the same attacks in future that might happen.

REFERENCE

[1]. C. Yin, Y. Zhu, J. Fei, and X. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," IEEE Access, vol. 5, pp. 21954–21961, 2017, doi: 10.1109/ACCESS.2017.2762418.

[2]. D. Agrawal, C. Agrawal, and H. Yadav, "A Machine Learning Based Intrusion Detection Framework Using KDDCUP 99 Dataset," vol. 4, no. 6, p. 11.

[3]. R. Taheri, M. Ahmadzadeh, and M. R. Kharazmi, "A New Approach For Feature Selection In Intrusion Detection System," p. 15.

[4]. A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A new feature selection model based on ID3 and bees algorithm for intrusion detection system," p. 8.

[5]. D. M. Abdulqader, A. M. Abdulazeez, and D. Q. Zeebaree, "Machine Learning Supervised Algorithms of Gene Selection: A Review," vol. 62, no. 03, p. 13, 2020. [9] R. Vijayanand, D. Devaraj, and B. Kannapiran, "A Novel Deep Learning Based Intrusion Detection System for Smart Meter Communication Network," in 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Tamilnadu, India, Apr. 2019, pp.1–3.doi:10.1109/INCOS45849.2019.8951344.

[6]. A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," Expert Systems with Applications, vol. 42, no. 5, pp. 2670–2679, Apr. 2015, doi: 10.1016/j.eswa.2014.11.009.

[7]. G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," Expert Systems with Applications, vol. 41, no. 4, pp. 1690–1700, Mar. 2014, doi: 10.1016/j.eswa.2013.08.066.

[8]. F. A. M. Bargarai, A. M. Abdulazeez, V. M. Tiryaki, and D. Q. Zeebaree, "Management of Wireless Communication Systems Using Artificial Intelligence-Based Software Defined Radio," Int. J. Interact. Mob. Technol., vol. 14, no. 13, p. 107, Aug. 2020, doi: 10.3991/ijim.v14i13.14211.

[9]. D. M. Manimekalai and G. Anupriya, "A Novel Intrusion Detection System using Data Mining Techniques," vol. 6, no. 6, p. 5, 2019.

[10]. N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," IEEE Trans. Emerg. Top. Comput. Intell., vol. 2, no. 1, pp. 41–50, Feb. 2018, doi: 10.1109/TETCI.2017.2772792.

[11]. Y. Xiao, C. Xing, T. Zhang, and Z. Zhao, "An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks," IEEE Access, vol. 7, pp. 42210–42219, 2019, doi: 10.1109/ACCESS.2019.2904620.

[12]. H. Yao, D. Fu, P. Zhang, M. Li, and Y. Liu, "MSML: A Novel Multilevel Semi-Supervised Machine Learning Framework for Intrusion Detection System," IEEE Internet Things J., vol. 6, no. 2, pp. 1949–1959, Apr. 2019, doi: 10.1109/JIOT.2018.2873125.

[13]. R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. AlNemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," IEEE Access, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[14]. G. Andresini, A. Appice, N. D. Mauro, C. Loglisci, and D. Malerba, "Multi-Channel Deep Feature Learning for Intrusion Detection," IEEE Access, vol. 8, pp. 53346–53359, 2020, doi: [10.1109/ACCESS.2020.2980937.

[15]. J. Gu and S. Lu, "An effective intrusion detection approach using SVM with naïve Bayes feature embedding," Computers & Security, vol. 103, p. 102158, Apr. 2021, doi: 10.1016/j.cose.2020.102158.

[16]. A. M. Abdulazeez and F. S. Khamo, "A Proposed Data Security Algorithm Based on Cipher Feedback Mode and its Simulink Implementation," vol. 4, no. 9, p. 10, 2013. [22] H. Wang, J. Gu, and S. Wang, "An effective intrusion detection framework based on SVM with feature augmentation," Knowledge-Based Systems, vol. 136, pp. 130–139, Nov. 2017, doi: 10.1016/j.knosys.2017.09.014.

[17]. N. Gao, L. Gao, Q. Gao, and H. Wang, "An Intrusion Detection Model Based on Deep Belief Networks," p. 6. [24] A. Pal Singh and M. Deep Singh, "Analysis of Host-Based and NetworkBased Intrusion Detection System," IJCNIS, vol. 6, no. 8, pp. 41–47, Jul. 2014.

[18]. R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. AlNemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," IEEE Access, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[19]. M. E. Aminanto and K. Kim, "Deep Learning in Intrusion Detection System: An Overview," p. 12.

[20]. . Alrawashdeh and C. Purdy, "Toward an Online Anomaly Intrusion Detection System Based on Deep Learning," in 2016 15th IEEE International Conference on Machine Learning and

Applications (ICMLA), Anaheim, CA, USA, Dec. 2016, pp. 195–200. doi: 10.1109/ICMLA.2016.0040.

[21]. S. KishorWagh, V. K. Pachghare, and S. R. Kolhe, "Survey on Intrusion Detection System using Machine Learning Techniques," IJCA, vol. 78, no. 16, pp. 30–37, Sep. 2013, doi: 10.5120/13608-1412.

[22]. D. Q. Zeebaree, A. M. Abdulazeez, D. A. Zebari, H. Haron, and H. N. A. Hamed, "Multi-Level Fusion in Ultrasound for Cancer Detection Based on Uniform LBP Features," p. 21, 2021.

[23]. K. A. Taher, B. Mohammed Yasin Jisan, and Md. M. Rahman, "Network Intrusion Detection using Supervised Machine Learning Technique with Feature Selection," in 2019 International Conference on Robotics,Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, Jan. 2019, pp. 643–646. doi: 10.1109/ICREST.2019.8644161.

[24]. S. M. Sohi, J.-P. Seifert, and F. Ganji, "RNNIDS: Enhancing network intrusion detection systems through deep learning," Computers & Security, vol. 102, p. 102151, Mar. 2021, doi: 10.1016/j.cose.2020.102151.

[25]. C. Kalimuthan and J. Arokia Renjit, "Review on intrusion detection using feature selection with machine learning techniques," Materials Today: Proceedings, vol. 33, pp. 3794–3802, 2020, doi: 10.1016/j.matpr.2020.06.218.

[26]. Ahmed, F. Y., Sreejith, R., & Abdullah, M. I. (2021, April). Enhancement of E-Commerce Database System During the COVID-19 Pandemic. In 2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE) (pp. 174-179). IEEE.

[27]. Alkawaz, M. H., Segar, S. D., & Ali, I. R. (2020). A Research on the Perception and use of Electronic Books Among it Students in Management & Science University. In 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA) (pp. 5256). IEEE.

[28]. Alkawaz, M. H., Rajandran, H., & Abdullah, M. I. (2020). The Impact of Current Relation between Facebook Utilization and E-Stalking towards Users Privacy. In 2020 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS) (pp. 141-147). IEEE.

BIOGRAPHIES

SURYA RAJ R is an Undergraduate student, department of computer science and engineering in paavai college of engineering.

SURENDER U is an Undergraduate student, department of computer science and engineering in paavai college of engineering.

SIVA KUMAR S is an Undergraduate student, department of computer science and engineering in paavai college of engineering.

Mr.R.MURUGESAN ,M.E,(Ph.D)is an Assistant Professor, Head of the department,Department of Computer Science and Engineering in Paavai College of Engineering.