e-ISSN 2320 –7876 www.ijfans.org
Vol.11, S Iss.1, 2022
Research Paper © 2012 IJFANS. All Rights Reserved, **UGC CARE Listed (Group -I) Journal**

INTRUSION DETECTION AND CRIME PREDICTION WITH MACHINE LEARNING AND IoT

Anirudh Kumar Tiwari[1], Bhavana Narain[2]

MATS School of IT, MATS University, Raipur, Chhattisgarh, India
Email ID: tiwarianirudh646@gmail.com,narainbhawna@gmail.com

Abstract

Crime prediction is a foremost requirement of the current time. Technology helps to process crime prediction in advance form. In this era of digitalization crime investigation and prediction are a top and foremost necessity. An action or commission which constitutes an offense and is punishable by Law is called a crime in today's time, crime has increased a lot and if the crime is identified at the right time and informed to the police or the government, then IoT and computer applications like IDS are a blessing of digital technology and are used to find criminal information, and it helps the police to get data. The purpose of our work is to design a prototype that helps the police in detecting crime locations. We have collected data set from IoT-based sensor. Data collected is pre-processed and arranged in an excel sheet. This data set is used in an ensemble classifier of machine learning to predict the crime. Ensemble classifier is an improved technique of classification which support azure machine learning. It gives high accuracy and avoids overfitting and classification. It helps improve machine learning results by comminuting several models. It creates multiple data set and at last various classifier to give an accurate result. When we use an ensemble classifier in our work, we get high accuracy if different base models misclassify different training. Ex- even if the base classifier accuracy is slow.Crime can be performed by an individual or group. It can commit against the government or private sector it may harm someone's reputation, physical harm or mental harm crime can cause direct harm or indirect harm to whomever the victim is. We have taken a condition that if any person is going somewhere and after seeing an accident when the photo of that accident is taken then automatically it will be sent to nearest police Station. For this, it is necessary to havean application designed by us both for the sender and the receiver. This whole matter will directly connect the police with the crime location which eases the police can reach that location. GPS will be used for location detection. In our work, we have collecteddatasets with the help of a digital camera that is attached to an IoT device. In the first partof our paper, we have discussed the grounds of our work under the introduction of crime,digital image processing, GPS, and IoT. In the second part of our work, we have discussedthe

Research Paper © 2012 IJFANS. All Rights Reserved,

methodology of our work here sensor board, and GPS setting has been discussed alongwith the dataset. There is a number of data collection technologies in the IoT. The most widely used technology is the Wireless sensor network (WSN) uses multi-hopping and self-organization to maintain control over the communication nodes.

1. Introduction

In this era of computer, IoT is a network that exchanges information by connecting many types of network objects and IoT uses IoT sensor for this process [1][2][3]. IoT is a concept with the help of which all work is easily possible and the main task of IoT is to design a better network by recognizing all the devices connected to the Internet. If a device has the facility of on / off, then we can call it part of IoT, During the last decades, information technologies supported the pc networks play a significant role in various spheres of the act. Information has become the organization's most precious asset. Network Intrusion Detection System (NIDS) can define the embedded process in networking for devices like Smart Sensor inspired devices & under a Service Oriented Architecture (SOA) to regulate independently as an anomaly-based NIDS or integrated, transparently, during a very Distributed Intrusion Detection System (DIDS) [4]. The quality datasets contain inconsistencies that degrade the performance and increase the computational time. This motivates us to pre-process the datasets and forms an efficient version that helps researchers to make a robust and price-effective model for intrusion detection [5]. A ramification of intrusion detection approaches is present to resolve this severe issue but the foremost problem is performance. It is vital to increase detection rates and reduce warning rates within the world of intrusion detection. Soon detect the intrusion, various approaches are developed [6]. An intrusion Detection System (IDS) can discover the malicious activities and irregularities within the network and provide an awfully important basis for network defense. thanks to the event of cloud computing, social networks, additionally as mobile cloud computing, IDS has become even more important than before ([7][8]). Through the following hyperparameter selection methods and KDD Cup 99 dataset chosen optimal network parameters and network topologies for DNNs [9]. Serious network attacks may cause damage to computer systems, network paralysis, dataloss, or leakage. Network intrusion detection systems (IDS) try and identify unauthorized, illicit, and anomalous behavior based solely on network traffic to support deciding in-network preventive actions by network administrators [10]. Many researchers are contributing to this field for the last twenty years. During this paper, certain literature survey has been done to display some past research work.

2. Networking Attacks

These networking attacks are an overview of the four major categories ([11][12][13]).

Denial of Service (DoS)

attack renders system resources unavailable to legitimate users. The attack during which the hacker makes a computing or memory resources too busy or too full to serve

Research Paper © 2012 IJFANS. All Rights Reserved

legitimate networking requests and hence denying users access to a machine e.g. apache, smurf, Neptune, ping of death, back, mail bomb, UDP storm, etc. are all DoS attacks.

Remote to User Attacks (R2L)

the attacker first succeeds is gaining a foothold on the remote system within the sort of user session. a far off to user attack is an attack during which a user sends packets to a machine over the online, which s/he doesn't have access to so on show the machines vulnerabilities and exploit privileges which a section user would wear the pc e.g., xlock, guest, xnsnoop, phf, sendmail dictionary, etc.

User to Root Attacks (U2R)

These attacks are exploitations during which the hacker starts off on the system with a typical user account and attempts to abuse vulnerabilities within the system so on realize super user privileges e.g., perl, xterm.

Probing

it's new attack threat for IDS. Probing is an attack within which the hacker scans a machine or a networking device so on determine weaknesses or vulnerabilities which can later be exploited so on compromise the system. this method is sometimes employed in processing e.g., saint, portsweep, mscan, nmap et.

3. Intrusion Detection System

Intrusion detection is defined because the strategy of intelligently monitoring the events occurring in an exceedingly very automatic processing system or network, analyzing them for signs of violations of security policy [14]. The primary aim of Intrusion Detection System (IDS) is to shield the obtainability, self-assured, and probity of critical networked information systems [15].

Research Paper © 2012 IJFANS. All Rights Reserved, **UGC CARE Listed (Group -I) Journal**

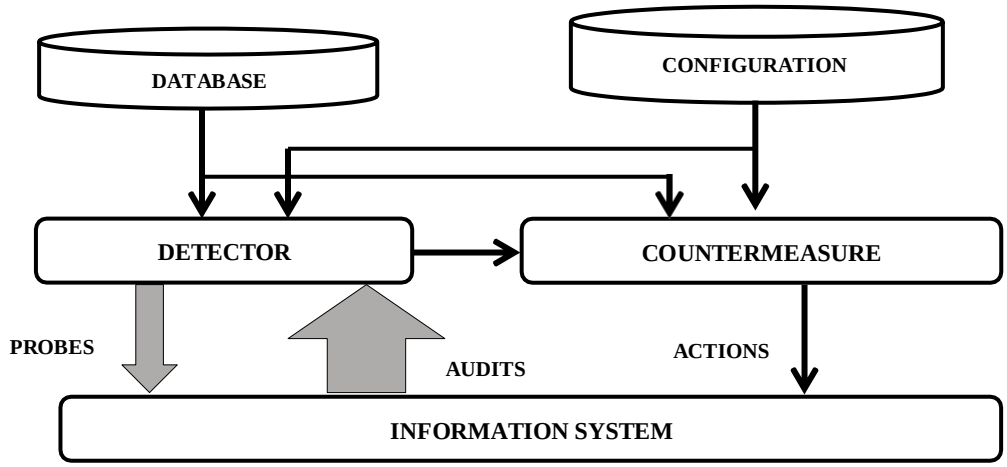


Fig.1: Intrusion Detection System

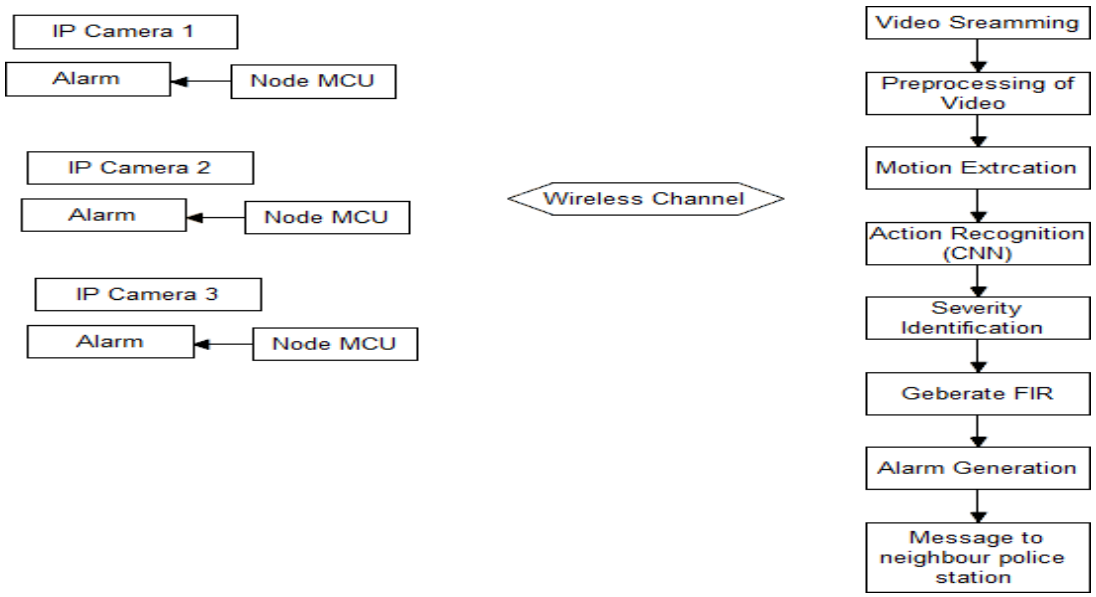


Fig. 2: Flow Diagram of Work Model

Attacks targeted by intrusion detection systems could even be divided into three forms user-level, system-level, and network-level. Typical user-level attacks include masquerade attacks [16]. The IDS researchers traditionally use three styles of metrics to check the performance in the network: False-Positive Rate (FPR), False-Negative Rate (FNR), and its complement, True Positive Rate (TPR) [17]. Recently, to safeguard computer networks and beat network security issues, Machine Learning (ML) approaches are implementedwithin the SDN-based Network Intrusion Detection Systems (NIDS). A

Research Paper © 2012 IJFANS. All Rights Reserved, **UGC CARE Listed (Group -I) Journal**

stream of advanced machine learning approaches the deep learning technology (DL) commences emerging within the SDN context ([18][19][20]).

The role of uncovering is to delete or remove the unneeded information from the audit trail. he increasing network output challenges during this NIDS system it define the possess compatible high and better performance processing and supply Rule-Based High-Performance Network Intrusion Detection System (RHPNIDS) for high-speed network which improves the system performance supported advanced approaches to both data collection and data analysis ([21] [22].

4. Classification of the Intrusion detection system

This research builds upon their work and introduces the very detailed and deep networking technique that these referenced research works don't describe [23]. There are two varieties of IDS: One is Network-based IDS and also the other is Host-based IDS. The NIDS monitors the data packets which is sending the networking from the network and the HIDS analyses the audit data of the operation system. [24]

Network-Based Intrusion Detection Systems (NIDS)

A network-based Intrusion Detection system detects malicious traffic on a network. Network Intrusion Detection Systems usually require promiscuous network access so on investigate all traffic including all unicast traffic [25]. Network Intrusion Detection Systems are passive devices that don't interfere. The Network Intrusion Detection System sniffs the net interface of the firewall in read-only mode & sends alerts toa Network Intrusion Detection System management several via a precise (read/write) network interface. [26]

Host Based IDS (HIDS)

Host Intrusion Detection System is an application that's accustomed check or monitoring a computer or network for suspicious activity, which might include intrusion by external actors like through the devices yet as misuse of resources or data by internal ones [27]. A bunch-based Intrusion Detection system is capable of the dynamic behavior andalso the state of an ADPS supported by how it's configured. The principal operation of host Intrusion Detection Systems relies upon the actual incontrovertible fact that successful intruders (hackers) will generally leave a trace of their activity.

Anomaly Based Detection

NIDS is the only way of defending against network-based attacks geared toward computer systems. These systems are utilized in most large-scale IT infrastructures. There are two main kinds of intrusion detection systems (a) signature-based and (b) anomaly-based. ABS systems build a statistical model like a mathematical model for describing the traditional network traffic and find any abnormal behavior that deviates from the model isidentified.

Research Paper © 2012 IJFANS. All Rights Reserved

Signature Based Detection

SBS systems rely on pattern recognition techniques where they are accustomed to maintaining the database which suggests a collection of information of signatures of previouslyknown attacks or fires and compare them with analyzed data after the analysis process. An alarm is raised when the signatures are matched

.

5. METHODOLOGYOGY

An action which constitutes an offence and is punishable by rule is called crime. It can be performed by individual or groups .it can commit against government or private sector. it may be harm someone reputation, physical harm or mental harm.

• Hacking: - it is illegal practice that harm some one's important data.

• Unwanted mass surveillance: - it is for personal interest it is considered as crime.

• Copyright infringement: - if someone infringes some one's protected copyright without permission and publishes then it is copyright infringement.

• Cyber extortion: - when someone hack email server, or computer system and demand for money then it is called cyber extortion.

• Cyber terrorism: - when someone hack government security then it is called cyber terrorism. For secure data we need to pay attention about

i) Security architecture

ii) Network diagram

iii) Risk management policy

iv) Security system

v) Disaster recovery plan

vi) Backup and restore procedure.

So that our main focus or our target is to reduce communication overhead & energy connection while maintaining high rate of service. We have collected data using IoT in Arduino board. This was developed by us. Data from 2017 to 2020 was collected through this method. Ensemble Learning Techniques was used to get the accuracy of two classifiers.

6. EXPERIENTIAL WORK

6.1 Ensemble Learning Technique: -

Under this technique we have used Random Forest and Bagging Classifier. We have compared the outcome of two with other classifiers.

Research Paper © 2012 IJFANS. All Rights Reserved

6.2 Random Forest: -

The Random Forest algorithm is a technique which creates a forest with the number of trees in decision tree. It can be used for degenerate, classification and other tasks. During training. Using Random Forest, the variable scan be ranked on the basis of their priority-based system.

Random Forest algorithm

BriefdefinitionofRandomForestalgorithmworksasalargecollectionofthecorrelateddecisiontreesWell, the name forest is because we use a lot of decision trees use them to make a classification.

Let's see

$$M=\begin{matrix} f_{A1} & f_{B1} & f_{C1} \\ f_{A2} & f_{B2} & f_{C3} \\ f_{An} & f_{Bn} & f_{n} \end{matrix}$$

We show here in this example a matrix M Suppose this matrix S is a matrix of training samples. In this case f A 1, f B 1, f C 1, these are a lot of features, for example the f A 1 is the feature A of the first sample And we can continue with all the given samples so that is up to N So, the f B N is the feature B of the Nth sample And also we have in the last column here the C 1 and C N, which means that we have lots of features and we have a training class So that the aim is to create a random forest to classify this sample set How the algorithm works.

6.3 Confusion matrix

A confusion matrix is a matrix that evaluates the performance of classification-based model.

Table1: Table Classification Accuracy of random forest Classifier (by Using the Data set with Pre-processing)

Year	Accuracy	TP Rate	FPRate	Precision	Recall	FMeasure
2017	0	0.	0	0. 356	0	.882
2018	88.333	.883	.051	.892	.883	.883
2019	98.963	.983	.006	.984	.983	..983
2020	97.444	.978	.009	.978	.978	.978

Research Paper © 2012 IJFANS. All Rights Reserved, **UGC CARE Listed (Group -I) Journal**

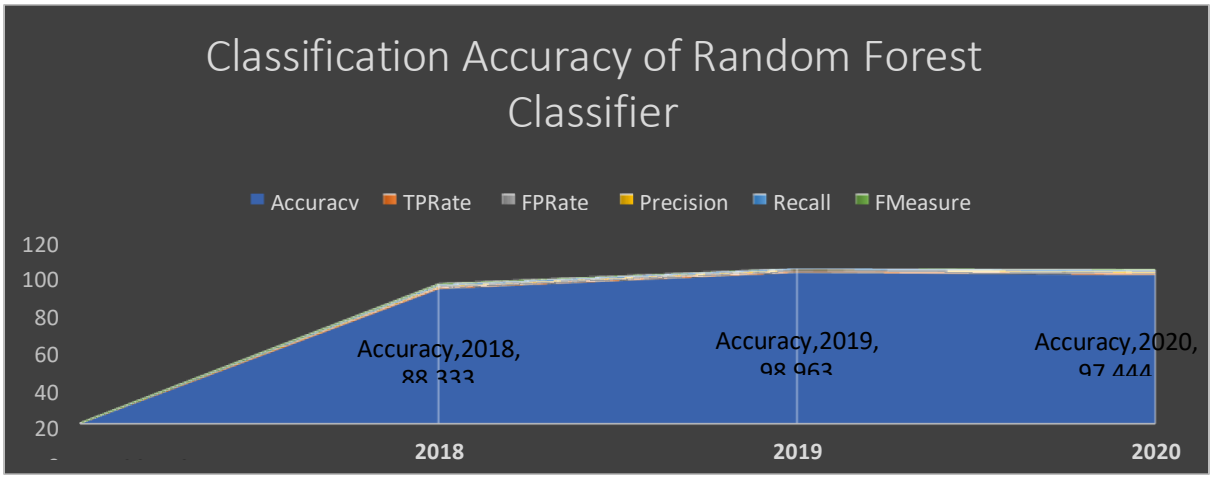


Figure3: Classification Accuracy of Random Forest Classifier (by Using the Data set with Pre-processing)

Confusion matrix

A confusion matrix is a matrix based technique for summarizing the performance of a classification algorithm for different sets .

Table2: Classification Accuracy of Bagging Classifier (by Using the Data set with Pre-processing)

Year	Accuracy	TP Rate	FPRate	Precision	Recall	FMeasure
2001	16.667	0.167	0.35	0. 11	0.167	.132
2005	71.667	.717	.011	.721	.717	.714
2010	94.444	.944	.017	.952	.944	..943
2015	95.556	.956	.014	.957	.956	.956

7 Comparison Study

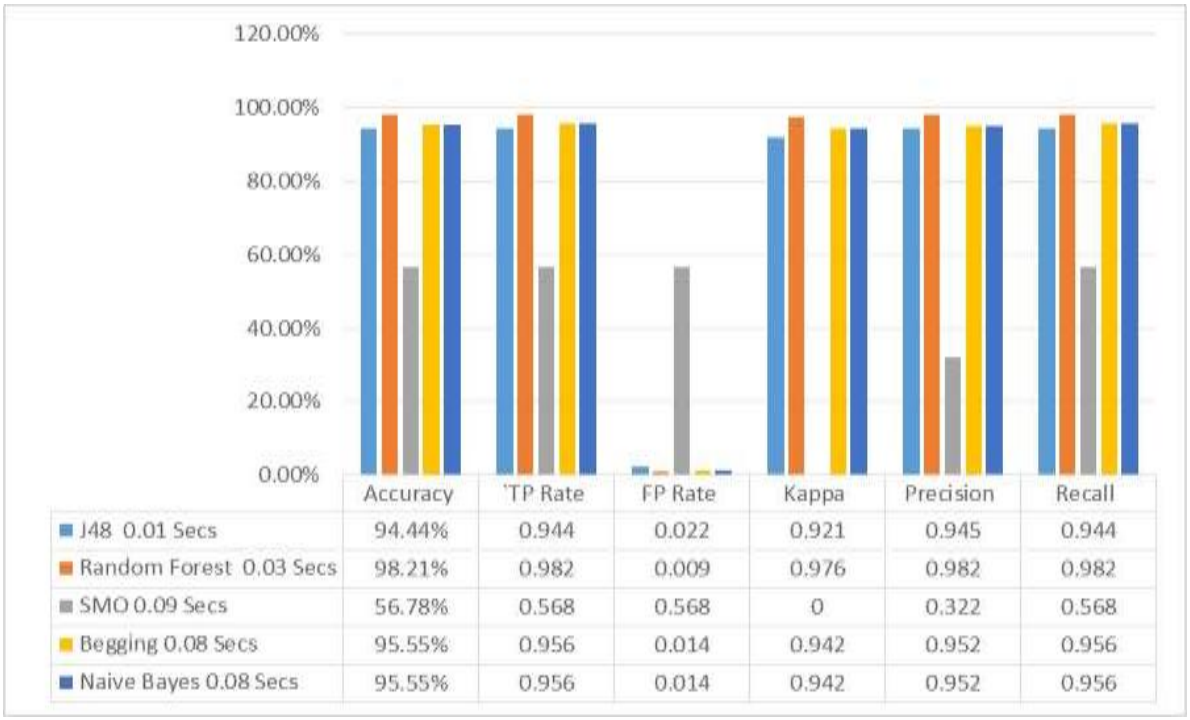
In the first phase, the sequential model is created. It is implemented using weka package and its results are compared with existing Machine Learning algorithms. Existing Machine Learning algorithms are implemented using Weka 3.9 library in java.

Research Paper © 2012 IJFANS. All Rights Reserved, **UGC CARE Listed (Group -I) Journal**

Comparison Of Different Classification Algorithm Performance on Different Outcomes

Table3: shows the classification of different outcomes acquired from the five-classifier utilized in this work while

Evaluation Metrics	**J48**	**Random Forest**	**SMO**	**Begging**	**Naive Bayes**
Time	0.01Secs	0.03Secs	0.09Secs	0.08Secs	0.08Secs
Accuracy	94.44%	98.21%	41.667%	95.55%	95.55%
`TPRate	0.944	0.982	0.417	0.956	0.956
FPRate	0.022	0.009	0.292	0.014	0.014
Kappa	0.921	0.976	0	0.942	0.942
Precision	0.945	0.982	0.322	0.952	0.952
Recall	0.944	0.982	0.417	0.956	0.956



Research Paper © 2012 IJFANS. All Rights Reserved, **UGC CARE Listed (Group -I) Journal**

8 Conclusion:

In this paper, we have presented literature survey on network intrusion detection system. Network security is playing vital role altogether styles of networks. Intrusion detection has attracted considerably more interest from researchers and industries so currently it's sensible choice for networking user for security purpose. After some years of research, the community still faces the matter of building reliable and efficient NIDS, which are capable of handling large amounts of knowledge, with changing patterns in real time situations. The scope of the work on classifying intrusion detection systems, reviewing the various methods of detecting anomaly, performance of those methods was supported past and up up to now works revealing the benefits and drawbacks of every of them.

References

[1] Francisco Macia-Perez, Francisco J. Mora-Gimeno, Diego Marcos-Jorquera," Network Intrusion Detection System Embedded on a Smart Sensor",IEEE,2010.

[2] Santosh Kumar Sahu, SauravranjanSarangi, Sanjaya Kumar Jena,"A Detail Analysis on Intrusion Detection Dataset",IEEE,©2014,978-1-4799-2572-8/14.

[3] Sravan Kumar Jonna lagadda, Ravi Prakash Reddy,"A Literature Survey and Comprehensive Study of Intrusion Detection",International Journal of Computer Applications,November2013, 0975–8887,Vol.81,No.16.

[4] Kai Peng, Victor C .M .Leung ,and Qingjia Huang,"Clustering Approach Based on Mini Batch Kmeans for Intrusion Detection System Over BigData",Special Section on Cyber-Physical-Social Computing and Networking, February 28, 2018, Digital Object Identifier 10.1109/ ACCESS.2018.2810267,Vol.6,2169-3536,2018IEEE.

[5] R. Vinaya kumar, Mamoun Alazab, K. P. Soman, Prabaharan Poornachandran, Ameer Al- Nemrat, and Sitalakshmi Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System",IEEEAccess,April11,2019,10.1109/ACCESS.2019.2895334.Vol.7,2169- 3536, 2019IEEE.

[6] Longzhi Yang, Jie Li, Gerhard Fehringer, Phoebe Barraclough, Graham Sexton, Yi Cao, "Intrusion Detection System by Fuzzy Interpolation",2017 IEEE international conferenceon fuzzy systems (FUZZ-IEEE), 2017,pp.1-6.

[7] Biswanath Mukharjee, L. Todd Heberlein and Karl N. Levitt, "Network Intrusion Detection", IEEE Network, May/June1994,0890-8044/94/$0.4.00,©1994.

[8] Clive Grace, "Understanding Intrusion Detection Systems", PC Network Advisor, www.itp- journals.com, September 2000, Issue 122, pp.11.

[9] Ahmed Awad E. Ahmed and Issa Traore,"Anomaly Intrusion Detection based on Biometrics", 2005 IEEE Workshopon Information Assurance United States Military Academy, WestPoint, NY, June2005,ISBN555555555,©2005IEEE.

[10] Robert MitchellandIng- Ray Chen,"A Survey of Intrusion Detection Techniques for Cyber- Physical Systems",ACM Computing Surveys, March 2014, Vol.46,No.4.

[11] Nasrin Sultana, Naveen Chilamkurti, Wei Peng, Rabei Alhadad, "Survey on SDN based network intrusion detection system using machine learning approaches", Peer-to-Peer Networking and Applications, https://doi.org/10.1007/s12083-017-0630-0, 12 January 2018.

[12] Christos Xenakis, Christoforos Panos, Ioannis Stavrakakis, "A Comparative Evaluation of Intrusion Detection Architectures for Mobile Ad Hoc Networks", computer & security 30(1), 63- 80, 2011.

[13] Abdelouahid Derhab, Abdelghani Bouras, Mustapha RedaSenouci, and Muhammad Imran, "Fortifying Intrusion Detection Systems in Dynamic Ad Hoc and Wireless Sensor Networks", Hindawi Publishing Corporation International Journal of Distributed Sensor Networks, December2014,ArticleID608162,pp.15,http://dx.doi.org/10.1155/2014/608162.

[14] kirtichoudhary, KomalDhing, ShivaniPacharne, SweetySamanta, Anujaphapale,"Hybrid Approach to wards IDSIPSIR Susing Reinforcement Learning", IJRECE, APRIL-JUNE2019, Vol.7, ISSUE:2,ISSN:2393-9028(PRINT)|ISSN:2348-2281(ONLINE)

[15] Quamar Niyaz, Weiqing Sun, Ahmad YJavaid, and Mansoor Alam, "A Deep Learning Approach for Network Intrusion Detection System", BICT2015, December 03-05, New York City, United States, Copyright©2016ICST, DOI10.4108/eai.3-12-2015.2262516.

[16] Chirag N. Modia, Dhiren R. Patela, Avi Patelb, Muttukrishnan Rajarajanb,"Integrating Signature Aprioribased Network Intrusion Detection System (NIDS) in Cloud Computing", 2nd International Conferenceon Communication, Computing & Security (ICCCS-2012),6,2012,905– 912.

[17] Hu Zhengbing, Li Zhitang,WuJunqi, "A Novel Network Intrusion Detection System (NIDS) Based on Signatures Search of DataMining", e-Forensics2008,January21-23,2008,Adelaide, Australia,©2008ICST978-963-9799-19-6.

[18] Jihyun Kim, Jaehyun Kim, Huong Le Thi Thu, and Howon Kim,"Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection", International Conference on PlatformTechnologyandService,Deccember31,2015

[19] V.Jyothsna, V. V. Rama Prasad, "A Review of Anomalybased Intrusion Detection Systems", International Journal of Computer Applications,

Research Paper © 2012 IJFANS. All Rights Reserved,

September2011,0975–8887,Vol.28,No.7.

[20] Abhishek Das, David Nguyen, Joseph Zambreno, Gokhan Memik ,and Alok Choudhary, "An FPGA-Based Network Intrusion Detection Architecture",IEEE Transactions on Information Forensics and Security, March2008,Vol.3,NO.1.

*[21]*Chuanlong Yin, Yuefei Zhu, JinlongFei, XinzhengHe ,"A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks", IEEEA ccess,2169-3536,2017IEEE,Vol.5,2017.

[22] Hung-Jen Liao, Chun-Hung RichardLin,Ying- ChihLin ,Kuang-YuanTung, "Intrusion detections system A comprehensive ereview", Journal of Network and Computer Applications 36, Accepted11, September2012, Available online23, September2012,16–24.

[23] Mehdi Hosseinzadeh Aghdam, and Peyman Kabiri, "Feature Selection for Intrusion Detection System Using Ant Colony Optimization", International Journal of Network Security, May 2016, Vol.18, No.3,PP.420-432.

[24] RenHuiGong, Mohammad Zulkernine, Purang Abolmaesumi, "A Software Implementation of a Genetic Algorithm Based Approach to Network Intrusion Detection", ACIS International Workshop(IEEE),2005,0- 7695-2294-7/05$20.00,©2005.

*[25]*Wei Lu and IssaTraore, "Detecting New Forms of Network Intrusion using Genetic Programming", Computational Intelligence, 2004, Vol. 20, No. 3.

[26] B.M.Aslahi-Shahri, R.Rahmani, M. Chizari ,A. Maralani, M. Eslami, M. J. Golkar, A. Ebrahimi, "A hybrid method consisting of GA and SVM for intrusion detection system ",Neural Comput&Applic,30July2015,DOI10.1007/s00521-015-1964-2.

Research Paper © 2012 IJFANS. All Rights Reserved,

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, S Iss.1, 2022
Research Paper © 2012 IJFANS. All Rights Reserved, **UGC CARE Listed (Group -I) Journal**