iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** VI **Month of publication:** June 2017

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Network Intrusion Detection Using Machine Learning in Vanets: A Review

Priyanka Gulati[1], Kamal Gupta[2]
[1,2] Computer Science and Engineering Department, GNI, Mullana, Haryana, India

Abstract- Vehicular Ad-hoc Network (VANETs) aids the vehicles to form a self-organized network in the absence of centralized infrastructure. It is the constituent of MANET and supports intelligent transport system (ITS). Each node is free to move independently in an ad-hoc network. In absence of centralized architecture, security becomes the crucial aspect of VANET. Cryptographic techniques such as digital signatures and encryption are unable to deal with unknown attacks. In order to deal with this issue, there is need of other technique to "detect and notify" these unknown attacks i.e. intrusion detection. An Intrusion Detection System (IDS) is a set of software that monitors a single or a network of computers for malicious activities (attacks) aiming at stealing or tampering data or corrupting network protocol usual routing behavior. This paper focuses on possible security attacks in VANETs and introduced the concept of intrusion detection system to combat against these attacks. It also discusses machine learning techniques for security in VANETs.
Keywords-Intrusion Detection System (IDS); Security; Attacks; Machine Learnin; Data Mining

I. INTRODUCTION

Nowadays, as vehicles are increasing day by day on roads, driving becomes challenging and dangerous [1]. Every year lots of people lose their lives in road accidents so safety becomes the major issue in human life. The productivity of traffic environment is being affected due to these accidents [2]. To decrease the number of road accidents, there is a technology i.e. VANET acronym for Vehicular Ad-hoc Network emerged that helps the vehicles to communicate with each other. It uses three types of communication between vehicles: Vehicle to Vehicle communication (V2V), Vehicle to Infrastructure communication (V2I) and Infrastructure to Vehicle [3]. VANET has no fixed infrastructure and supports ITS (Intelligent Transport System). It is a part of MANET that considers the moving cars as nodes to create a self - organized network [4]. In VANET, communication among vehicles can be achieved through DSRC (Dedicated Short Range Communication) that uses IEEE 802.11p standard which is an accepted enhancement to the IEEE 802.11 standard to append WAVE (Wireless Access in Vehicular Environments) [1]. WAVE provides vast range of information to the drivers to enhance road safety. On-Board Unit (OBU) and Road Side Unit (RSU) are the WAVE constituents [5]. OBU is a wave device equipped in the vehicle that is organized for communication with the external network. RSU is a device which is fixed across the roads and helps to communicate between the moving vehicles and the fixed infrastructure. RSU helps to expand the communication range of ad-hoc network and deliver internet connectivity to OBUs [6].

Security is the crucial aspect in the VANETS. In order to provide security to the system, VANET should deal with the Authentication, Availability, Non-Repudiation, Integrity, and Confidentiality. In VANETS, due to high mobility, rapidly changing network topology, real time constraint, data consistency liability, environmental impact, vehicles are not stable [2]. So, security is still an issue in ad-hoc network. There are innumerable attacks that affect on the security of the VANET:

A. Denial of Service (DoS) Attack

In DoS attack, the attacker sends inappropriate messages on the communication channel to jam the network using single computer and single internet connection so that the network services make unavailable for the legitimate users [7].

B. Sybil Attac

In Sybil attack, the misbehaved vehicle sends various messages to other vehicles and every message contains the fake source ID so that misconception is provided to neighboring vehicle by sending false messages [8].

C. Distributed Denial of Service (DDoS) Attack

In DDoS attack, attackers fire attacks by sending multiple messages from distinct locations using multiple computers and internet connections to consume the bandwidth of targeted resources or to slower down the network [7].

©IJRASET: All Rights are Reserved

D. Timing Attack

This attack is very dangerous attack because of the misbehaved vehicle get the message but do not forward it to the neighboring vehicle on time to create a delay and hence collision is created between the vehicles [9].

E. Bogus Information Attack

The attacker's fire bogus information attack can be insider or outsider attack. In this attack, attackers launch attack by sending false information in the network to influence the performance of the drivers [10].

II. INTRUSION DETECTION SYSTEM

A fruitful approach to protect the network from various attacks is the implementation of Intrusion Detection System (IDS). Intrusion means to break in the system by compromising the integrity, confidentiality and availability of the resource. Intrusion Detection means to recognize the interference done by the attacker to maintain the security of the system. Intrusion Detection System is a system that checks the abnormal activities of an entity in the network. If it detect any malicious activity then alerts the system administrator. Two key attributes that need to be identified with respect to IDS are False alarm ratio and Detection ratio. For efficient IDS, False alarm ratio should be less and detection ratio should be high. Fig 1 gives detailed classification of IDS. There are basically two types of IDS:

A. Host Based IDS

Host Intrusion Detection System (HIDS) runs on a single host on the network and monitors the incoming and outgoing packets from the device. If it detects any unauthorized attempt then alert is sent to the administrator to examine the suspicious activity. In HIDS, anti-threat applications are installed on every computer to monitor the traffic on the network.

B. Network Based IDS

Network Intrusion Detection System (NIDS) examine the network traffic to diagnose the threats such as Denial of Service (DoS) attacks that break the security of the network. If any suspicious activity is identified by the system then alert is sent to the system administrator. In NIDS, anti-threat application is installed at the specific points such as server to invigilate the traffic on the network.

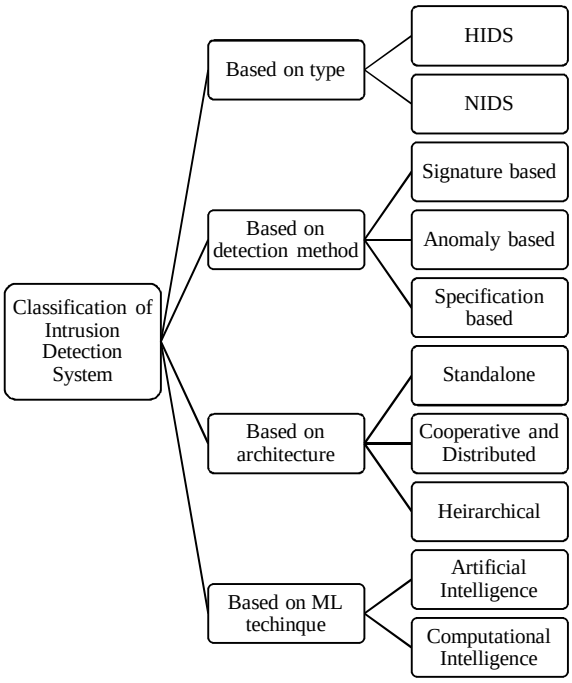


Fig 1: Classification of IDS

C. Based on detection method, IDS can be categorized as

©IJRASET: All Rights are Reserved

1) *Signature Based IDS:* This is also known as knowledge based detection or misuse detection. In Signature based ids, predefined series of pattern or signature of attacks store into the database and generates an alarm when the attack is done by the attacker matches with the pre-registered signature of attacks [11].
2) *Anomaly Based IDS:* This is also known as Behavior based detection. Anomaly based detection works on the normal behavior such as monitoring regular activities , kernel information, system logs etc that is stored into the database and generates an alarm if any suspicious activity is detected on the device that contradict the systems normal behavior.
3) *Specification Based IDS:* This is a hybrid approach which is a combination of Signature based IDS and Anomaly based IDS. Specification based IDS looks for the abnormal behavior at the system level and sets some parameters about the intrusion into the database that a program must satisfy and an attack is detected if program fails to meet the conditions stored into the database [12-13].

D. *On the basis of architecture, IDS can be classified into three ways [14]*

1) *Standalone IDS:* In Standalone IDS, every node runs autonomous IDS and is capable to recognize intrusion. Therefore, no data is shared with the other node. Hence, it results that every node is able to run and execute an independent IDS
2) *Cooperative and Distributive IDS:* In this architecture, intrusion is detected by transferring the information or alerts to other nodes in the network. In addition, every node cooperates with the other node to detect attacks. The major drawback of this architecture is that the performance of the network decreases due to the exchange of information between the IDS agents.
3) *Heirarichal IDS:* In Heirarichal IDS, network is isolated into set of groups called clusters. Each cluster has its leading head. The cooperation is accomplished between the leading head and every member of the same cluster. An alert is send to the leading head if the member node is not able to detect an attack. So, leading head grant the permission to monitor the network if something seems to be malicious in his cluster. In this approach, performance of the network increases due to less cooperation between the leading head and its members.

III. MACHINE LEARNING TECHNIQUES FOR NETWORK INTRUSION DETECTION

Machine learning refers to the ability of a computer program to learn from a set of inputs either in a supervised (by being actively trained), or unsupervised (by exploring the characteristics of raw data on its own) fashion, in order to provide answers to questions that it wasn't specifically designed to know the answer to. Machine Learning based techniques for intrusion detection are classified into two categories: Artificial Intelligence (AI) techniques and Computational Intelligence (CI) techniques. AI techniques include classical methods such as statistical modeling while CI techniques include nature-inspired methods that are used to deal with complex problems which classical methods are unable to solve. *Computational Intelligence* integrates artificial neural networks, evolutionary computation, swarm intelligence, fuzzy logic and artificial immune systems. Conventional AI techniques try to mimic human intelligence through symbol manipulation and symbolically structured knowledge bases. This approach limits the applications to which conventional AI can be applied. On the other hand, CI deals with subsymbolic knowledge processing i.e. numerical knowledge representation and processing.

A. *Artificial Intelligence Techniques*

1) *Support Vector Machine (SVM):* SVM is a data mining technique used for classification and regression problems. Classification consists of training set and testing set to diagnose the attacks. SVM builds up the hyperplanes by mapping the pattern vectors to higher dimensional spaces. It uses the kernel function that transforms the linear model to non - linear model such as Polynomial and Gaussian. Bhavsar et al. [15] proposed an intrusion detection system using support vector machine (SVM) to detect intrusion into the network. SVM consists of classification and verification method to lower the extensive training time. To achieve this goal, it has used three functions: Gaussian Kernel (Radial Basis Function), Polynomial kernel and Sigmoid kernel. Chen et al. [16] proposed a SVM model based on compressive sampling for feauture compression for network anamoly detection.
2) *Pattern Matching (PM):* PM is a data mining technique and also known as Misuse detection or Signature Base detection. Pattern matching is the act of auditing the fixed array of bytes within an individual packet. It fires an alarm, if it determines suspicious in the network. Karthiga et al. [17] proposed a state traversal mechanism incorporate with finite state machine (FSM) to identify network intrusion with the use of optimized pattern matching algorithm and are also responsible for reduction of memory space required during the implementation of FSM. To achieve this goal, longest common substring algorithm is used and then the outcomes are correlated with the AC algorithm and the bit split algorithm.
3) *Decision Trees:* It predicts the value of a target variable based on several input variables. In decision tree model, interior nodes

©IJRASET: All Rights are Reserved

correspond to one of the input variables, leaves represent value of target variable given the values of the input variables by the path from the root to the leaf and branches represent logical conjunctions of features leading to particular value of target variables. Decision tree models are of two types: classification trees in which target variable takes a finite set of values and regression trees in which it takes continuos values. Rai et al. [18] proposed Decision Tree Split algorithm based on C4.5 decision tree. The feature selection is done using information gain and the split value is selected in such a way that makes the classifier unbiased towards most frequent values. The classifier used NSL-KDD (Network Security Laboratory Knowledge Discovery and Data Mining) dataset and accuracy achieved is compared with other existing classifiers such as Classification and Regression Tree (CART), Naïve Bayes (NB) Tree, and Alternating Decision (AD) Tree. Kumar et al. [19] compared the performance of different decision tree classifiers such as ID3, C4.5 and C5 algorithms for misuse and anomaly attack detection using KDD99 dataset and found that C5 outperformed all.

4) *Naive's Bayesian Classifier:* It is a statistical technique and a supervised learning algorithm based on Baye's Theorem i.e. P (A|B) = P (B|A) P (A)/P (B). It handles broad database and is used to detect malicious network activity. Altwaijry et al. [20] presented a Bayesian based intrusion detection system to detect the feasible intrusions over the network. All attack records and normal records are tested using Bayesian filter and then outcomes are compared with the following algorithms: Gaussian classifier, SVM, K-means clustering etc. to increase the detection rate of R2L attacks.

5) *K-means clustering:* K-Means clustering partitions *n* different objects into *k* clusters based on distance function such that each object belongs to the cluster with the nearest mean. Brar and Sharma [21] proposed a novel density based k-means clustering (DBKmeans) model of IDS. The proposed model works on the KDD99 dataset and makes specific patterns of network attacks and normal queries by labeling the dataset. The results obtained clearly revealed improved accuracy and detection rate with lower false positive rate.

B. Computation Intelligence Techniques

1) *Artificial Neural Network (ANNs):* ANNs are inspirited by biological neural networks. It presents a system that is hooked by neurons which is responsible for transferring the messages. ANNs are used for the pattern recognition or to detect the well known attacks. Ting [22] proposed IDS based on the principle of Back propagation (BP) neural network to solve the efficiency problems such as slow training process and slow detection. The proposed BP algorithm and IDS was implemented using four modules: packet capturing, data analysis and processing, neural network and alarm generator to detect the malicious activity in the network. Al-Jarrah et al. [23] presented a smart system to analyze the attack behavior using TDNN neural network. This smart system used Principle component neural network to analyze attacks and a classification module to determine the port scan attacks. Principle components are extracted using Generalized Hebbian Algorithm (GHA). Liang Hu et al. [24] presented a feature selection algorithm and three neural network algorithms to enhance the development of IDS. Back Propagation (BP), Radial Basis Function (RBF) and Neural Networks with Random Weights (RNN) were the three algorithms used to figure out the feasibility of the IDS combined with the feature selection algorithm.

2) *Fuzzy Logic:* It is a form of multiple-valued logic which is used to determine the severity of maliciousness of the node. Fuzzy logic can be enforced in hardware, software or both. Fuzzy logic provides precise results based upon the noise in the input. It follows a simple rule based approach i.e. If X and Y then Z to determine the attacks or intrusions. This system is also known as Fuzzy Inference System (FIS). Harikishan et al. [25] proposed a novel approach called IDS using Fuzzy inference system to encounter the intrusion behavior within the network. Using Sugeno Fuzzy Inference approach and ANFIS editor, an accurate attack was detected. Naik [26] detected a port scan attack using Fuzzy inference based IDS i.e. FI Snort to anticipate the level of threat. It consists of four components. The first component is FI-Snort component i.e. a combination of fuzzy inference and Snort monitored and analyzed the traffic data. The second component collects the data and captured the various parameters using Snort. The third component performed data analysis using Wire shark networking tool and Snort. The fourth component took the input variables and converted those into the output variables to detect the level of port scan attack.

3) *Evolutionary Computation:* Evolutionary Computation or Genetic Algorithm is a technique that imitates biological evolution as a perceptive approach which is based on Darwin's ideology of evolution -"Survival of fitness". It is used to detect network anomalies. Following are the factors that will have the critical impact on the efficiency of the algorithm: Fitness function, Population, GA Parameters i.e. initialization, selection, crossover, mutation. Hoque et al. [27] implemented a novel approach of intrusion detection system (IDS) using GA which uses evolutionary theory in order to refine the traffic data and to lessen the complexity. The proposed Genetic Algorithm guarantees to improve the fidelity problem, resource usage problem, reliable

©IJRASET: All Rights are Reserved

problem that are found in the existing systems. Benaicha et al. [28] presented a GA which recognizes the DoS attacks and combined with the IDS which are responsible to increase the network intrusion detection rate and reduces the false positive rate. Genetic Algorithm optimizes the attack schemes in audit files. The proposed algorithm guarantees to improve the search time in audit data without lowering the performance of the system.

4) *Artificial Immune System (AIS):* AIS is a promising approach to deal with the Host Based Intrusion Detection System (HIDS). AIS are flexible systems inspirited by immunology and recognized immune functions, models which are enforced to problem solving. Negative Selection Algorithm, Immune Network Algorithms, Dendritic Cell Algorithms are used to determine the behavior of AIS. Rathee et al. [29] presented artificial immune system (AIS) based statistical model for identification of intrusion. It helps to detect various kinds of DoS attacks over the network. To identify the intrusion probability, statistical analysis based global approach is characterized which can be used on wireless network and to identify the probability of the attack over the network, an intelligent artificial immune system based approach is described.

5) *Swarm Intelligence:* The term Swarm Intelligence (SI) was coined by Beni and Wang [30] in 1989. The algorithms based on swarm intelligence draw their inspiration from the behavior of animals such as insects and birds, and their unique capability to solve difficult tasks in the form of swarms by socially interacting with each other and the environment. SI techniques aim at solving hard problems by the employing multiple agents without the aid of any supervision. Every agent roams in the search space and collaborates with others via direct or indirect communications for solving complex tasks such as finding classification rules for misuse detection and clusters for anomaly detection and achieving an optimal solution. Chung and Wahid [31] proposed an intrusion detection system by using intelligent dynamic swarm based rough set for feature selection of the network traffic and simplified swarm optimization for intrusion data classification. In addition, a weighted local search (WLS) strategy is incorporated in SSO to improve its performance. The performance of the proposed model is evaluated on KDDCup 99 dataset. Enache and Patriciu [32] proposed an IDS model based on Information Gain for feature selection and SVM as classifier. The parameters for SVM are further selected by a swarm intelligence algorithm. NSL-KDD data set was used for performance evaluation and authors show that proposed model achieved higher detection rate and lower false alarm rate as compared to SVM.

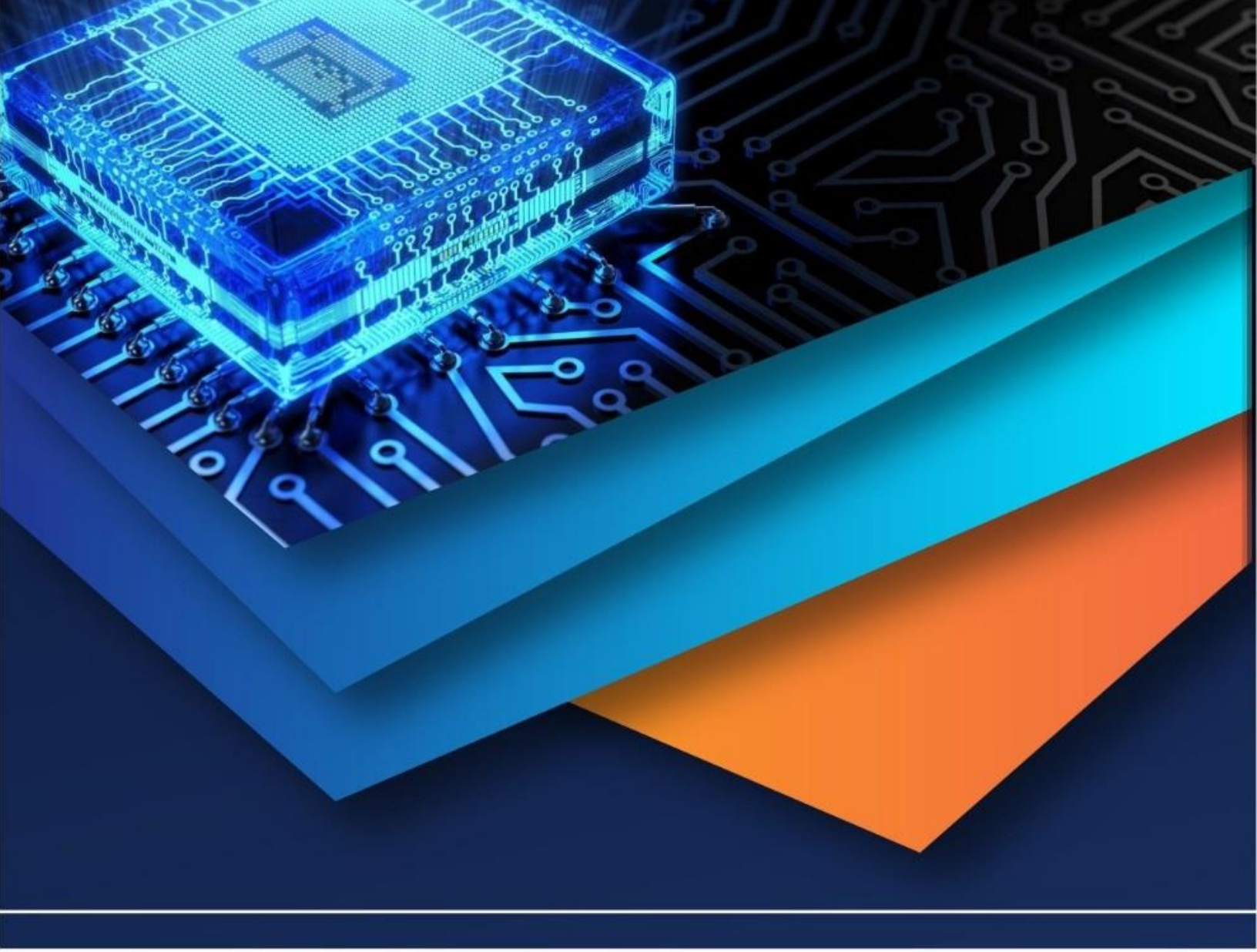
REFERENCES

[1] Patel, k. Shah, "a survey on vehicular ad hoc networks," iosr journal of computerengineering (iosr-jce), vol. 15, no. 4, pp. 34-42, 2013.
[2] M. K. N. S. R.s. Raw, "2. Raw, r. S., kumar, m., & singh, n. (2013). Security challenges, issues and their solutions for vanet., 5(5), 95.," international journal of network security & its applications, vol. 5, no. 5, pp. 95-105, 2013
[3] H. Hasrouny, a.e. Samhat, c. Bassil, and a. Laouiti, “vanet security challenges and solutions: a survey,” vehicular communications, vol. 7, pp. 7-20, 2017.
[4] G. Samara, w.a.h al-salihy, r. Sures, " security issues and challenges of vehicular ad hoc networks (vanet)," in 2010 4th international conference on new trends in information science and service science (niss), malaysia, 2010
[5] Y. J. Li, "an overview of the dsrc/wave technology," in springer berlin heidelberg in quality, reliability, security and robustness in heterogeneous networks, australia, 2015
[6] A. Singh, m. Singh, "a comprehensive review on vehicular ad hoc network," international journal of advanced research in computer and communication engineering, vol. 4, no. 4, pp. 462-468, 2015.
[7] I. A. Sumra, i. Ahmad, h. Hasbullah, j. L. B. A. Manan, "classes of attacks in vanet in electronics,, (pp. 1-5). Ieee.," in 2011 saudi international in electronics, communications and photonic conference (siecpc), riyadh, 2011.
[8] S. S. Tangade, s. S. Manvi, "a survey on attacks, security and trust management solutions in vanets," in 2013 fourth international conference in computing, communications and networking technologies (icccnt), banglore, 2013
[9] I.a. Sumra, j. L. Ab manan, h. Hasbullah, "timing attack in vehicular network," in world scientific and engineering academy and society (wseas) in proceedings of the 15th wseas international conference on computers, corfu island.
[10] H. La, a. Cavalli, "security attacks and solutions in vehicular ad hoc networks: a survey," international journal on adhoc networking systems (ijans), vol. 4, no. 2, pp. 1-20, 2014.
[11] H. J. Liao, c. H. R. Lin, y. C. Lin, k. Y. Tung, "intrusion detection system: a comprehensive review," journal of network and computer applications, vol. 36, no. 1, pp. 16-24, 2013
[12] S. Kaushik, s. Sharma, "securing ad hoc networks for inrusion detection, a study," international journal of innovations & advancement in computer science (ijiac), vol. 4, pp. 16-23, 2015
[13] R. Michtell, r. Chen, "a survey of intrusion detection in wireless network applications," computer communications, united states, 2014
[14] B. E. O. M. Erritali, "a survey on vanet intrusion detection systems," international journal of engineering & technology , vol. 5, no. 2, pp. 66-69, 2013
[15] Y. B. Bhavsar, k. C. Waghmare, "intrusion detection system using data mining technique: support vector machine," international journal of emerging technology and advanced engineering, vol. 3, no. 3, pp. 581-586, 2013.
[16] S. Chen, m. Peng, h. Xiong and x. Yu, x., 2016. Svm intrusion detection model based on compressed sampling. Journal of electrical and computer engineering, pp. 1-6, 2016
[17] K. Rai, m. S. Devi, and a. Guleria. "decision tree based algorithm for intrusion detection" international journal of advanced networking and applications, vol. 7, no. 4, pp. 2828-2834, 2016

©IJRASET: All Rights are Reserved

[18] M. Kumar, m. Hanumanthappa and t. V. S. Kumar, "intrusion detection system using decision tree algorithm," 2012 ieee 14th international conference on communication technology, chengdu, pp. 629-634, 2012
[19] H. Altwaijry, s. Algarny, "bayesian based intrusion detection system," journal of king saud university – computer and information sciences, vol. 24, no. 1, pp. 1-6, 2012
[20] R. Brar and n. Sharma. "a novel density based k-means clustering algorithm for intrusion detection," journal of network communications and emerging technologies (jncet), vol. 3, no. 3, pp. 17-22, 2015
[21] C. Ting, "detection system and the realization of the principle of bp neural network based intrusion," in 2015 seventh international conference on measuring technology and mechatronics automation (icmtma), nanchang, 2015.
[22] O. Al-jarrah, a. Arafat, "network intrusion detection system using neural network classification of attack behavior," journal of advances in information technology, vol. 6, no. 1, pp. 1-8, 2015.
[23] L. Hu, z. Zhang, h. Tang, n. Xie, "an improved intrusion detection framework based on artificial neural networks," in 2015 11th international conference on natural computation (icnc), zhangjiajie, 2015
[24] A.harikishan, p.srinivasulu , "intrusion detection system using fuzzy inference system," international journal of computer & organization trends, vol. 3, no. 8, pp. 345-352, 2013
[25] N. Naik, "fuzzy inference based intrusion detection system: fi-snort," in 2015 ieee conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing (cit/iucc/dasc/picom), liverpool, 2015
[26] M. S. Hoqque, m. A. Mulkit, m. A. N. Bikas, "an implementation of intrusion detection system using genetic algorithm," international journal of network security & its applications (ijnsa), vol. 4, no. 2, pp. 109-120, 2012
[27] S. E. Benaicha, l. Saoudi, s. E. B. Guermeche, o. Lounis, "intrusion detection system using genetic algorithm," in science and information conference (sai), london, 2014
[28] Y. Y. Chung and n. Wahid, "a hybrid network intrusion detection system using simplified swarm optimization (sso)," applied soft computing, vol. 12, no. 9, pp.3014-3022, 2012
[29] G. Beni and j. Wang, "swarm intelligence in cellular robotics systems", in: proceedings of nato advanced workshop on robots and biological system, pp. 703-712, 1989
[30] G. Rathee, p. Bano, s. Singh, "artificial immune system based statistical model for intrusion identification," international journal of computer science and mobile computing (ijcsmc), vol. 4, no. 6, pp. 170-176, 2015.
[31] A. C. Enache and v. V. Patriciu, "intrusions detection based on support vector machine optimized with swarm intelligence," 2014 ieee 9th ieee international symposium on applied computational intelligence and informatics (saci), timisoara, 2014, pp. 153-158.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)