# Intrusion Detection System Based on Advance K-Nearest Neighbour using Fitness Function

**B.Gayathri [1] Dr.Antony Selvadoss Thanamani[2]**
[1]Research Scholar [2]Associate Professor
[1,2]Department of Computer Science & Engineering
[1,2]NGM College Pollcahi

***Abstract***— *Network Intrusion Detection Systems are placed at a strategic point or points within the network to monitor traffic to and from all devices on the network. Ideally, you would scan all inbound and outbound traffic, however doing so might create a bottleneck that would impair the overall speed of the network. To implement these systems various researchers introduced numerous machine learning techniques like Decision Trees, Support Vector Machines, Artificial Neural Networks, Linear Genetic Programming, Genetic Algorithms, Fuzzy Inference Systems, Rule Based Approach and their ensemble approaches with the intent to predict the data either normal or abnormal. In this paper we propose genetic programming whereas the data mining classifier which performs the classification process is Advanced K-Nearest Neighbour. An experimental evaluation using real datasets shows that intrusion detection system calculation.*

**Keywords—** An intrusion detection system (IDS), Linear Genetic Programming, K-Nearest Neighbour classifier, Rule Based Approach

## I. INTRODUCTION

An Intrusion Detection System is an application used for monitoring the network and protecting it from the intruder. With the rapid progress in the internet based technology new application areas for computer network have emerged [7]. In instances, the fields like business, financial, industry, security and healthcare sectors the LAN and WAN applications have progressed. All of these application areas made the network an attractive target for the abuse and a big vulnerability for the community [7]. Malicious users or hackers use the organization's internal systems to collect information's and cause vulnerabilities like Software bugs, Lapse in administration, leaving systems to default configuration [8]. As the internet emerging into the society, new stuffs like viruses and worms are imported. The malignant so, the users use different techniques like cracking of password, detecting unencrypted text are used to cause vulnerabilities to the system. Hence, security is needed for the users to secure their system from the intruders. Firewall technique is one of the popular protection techniques and it is used to protect the private network from the public network. IDS are used in network related activities, medical applications, credit card frauds, Insurance agency [8]. The goal of intrusion detection is to monitor the network assets to detect anomalous behavior and misuse in network [12]. Intrusion detection concept was introduced in early 1980's after the evolution of internet with surveillance end monitoring the threat [10]. There was a sudden rise in reputation and incorporation in security infrastructure. Since then, several events in IDS technology have advanced intrusion detection to its current state [10]. James Anderson's wrote a paper for a government organization and imported an approach that audit trails contained important information that could be valuable in tracking misuse and understanding of user behavior [10]. Then the detection appeared and audit data and its importance led to terrific improvements in the subsystems of every operating system [12]. IDS and Host Based Intrusion Detection System (HIDS) were first defined. In 1983, SRI International and Dorothy Denning began working on a government project that launched a new effort into intrusion detection system development [8]. Around 1990s the revenues are generated and intrusion detection market has been raised. Real secure is an intrusion detection network developed by ISS. After a year, Cisco recognized the priority for network intrusion detection and purchased the Wheel Group for attaining the security solutions [8]. The government actions like Federal Intrusion Detection Networks (FID Net) were designed under Presidential Decision Directive 63 is also adding impulse to the IDS [8].

## II. EXISTING SYSTEM

In the existing system architecture which is the combination of host based and network based intrusion detection system. In particular IDS capture packets and call to detector agent where detector agent pass capture packets to rule matching process where rule matching process check attacks criteria from the database , where we have already defend and stored rule to find attack. This leads to degrade in performance. After completing this process alarm will activate if any type of attack find in the captured packet otherwise it will be deactivate and this processes will continue till on the existing system. So that packets may loss the data during the time of data transfer.

## III. IMPLEMENTATION

### *A. Genetic Programming*

Genetic programming is considered as an evolutionary technique which is becoming very popular for detecting intrusions. It was introduced by koza and its group. It is famous for its ability to learn relationship which are not visible in data and then finally present them in a mathematical manner. Its application can be found in designing classifier for a multiclass problem, generation of crips and fuzzy rules, classification of image, multicategory classification problem ect. The main motive of using this machine learning technique is to make the system capable of recognizing new information and new task patterns.GP has the potential to figure out the new attacks or novel data by learning from previous patterns. It assesses each randomly generated candidate on the basis of fitness function.

### *B. AKNN – Advanced K nearest Neighbour*

K-nearest neighbor (KNN) classification algorithm is a data mining algorithm which is theoretically mature with low complexity. The basic idea is that, in a sample space, if most of its nearest neighbor samples belong to a category, then the sample belongs to the same category. The nearest neighbor refers to the single or multidimensional feature vector that is used to describe the sample on the closest, and the closest criteria can be the Euclidean distance of the feature vector.The formula for calculating the distance between the two variables during the time of intrusion

$$d = \sqrt{\sum_{i=1}^{v} (p_{1i} - p_{2i})^2}$$

where the difference between two variables' values is taken, and squared, and summed for p persons (in our example p=3). Only one distance would be computed – between v1 and v2. Let's do the calculations for finding the Euclidean distances between the three persons, given their scores on two variables. Normalised Euclidean distance produces its "normalisation" by dividing each squared discrepancy between attributes or persons by the total number of squared discrepancies

## IV. METHODOLOGY

### *A. Advanced K-nearest neighbour (Knn) algorithm pseudocode:*

Let (Xi, Ci) where i = 1, 2……., n be data points. Xi denotes feature values & Ci denotes labels forXi for each i. Assuming the number of classes as 'c' Ci ∈ {1, 2, 3, ……, c} for all values of i.

Let x be a point for which label is not known, and we would like to find the label class using advanced k-nearest neighbour algorithm

### *B. Knn Algorithm Pseudocode:*

1) start
2) Upload dataset DT, and declare variable.
3) Calculate Euclidean distance between the variables.
4) Dr = data requirements and X , Xi are the variables to be accessed from the dataset
5) Calculate "d(x, $x_i$)" i =1, 2, ….., n; where d denotes the Euclidean distance between the points.
6) Arrange the calculated n Euclidean distances in non-decreasing order.
7) Let k be a +ve integer, take the first k distances from this sorted list.
8) Find those k-points corresponding to these k-distances.
9) Let $k_i$ denotes the number of points belonging to the $i^{th}$ class among k points i.e. $k \geq 0$
10) If $k_i > k_j \; \forall \; i \neq j$ then put x in class i.
11) From the requirements and Class i the intruder will be filtered.
12) Return class i
13) Stop.

## V. RESULT

Preprocess | Cancel | Available Data : 800

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| emp001 | bar|raj | m | ceo | level 4 | 23-02-15 | 255.100.162.240 | 36 | 255.209.197.126 | 73 | 252.120.186.83 | 213 | 3359931 | 9860231 | 8669426 | yes | |
| emp002 | ramdin verma | m | chief incharge | level 4 | 24-02-15 | 255.100.236.15 | 214 | 255.198.223.72 | 121 | 251.186.19.36 | 222 | 3313697 | 1190065 | 2431177 | yes | |
| emp003 | sharat chandran | m | team leader | level 3 | 11-03-15 | 255.100.162.240 | 184 | 255.203.177.198 | 49 | 252.244.74.20 | 61 | 4359854 | 2233429 | 8233591 | yes | |
| emp004 | birender mandal | m | team leader | level 3 | 19-03-15 | 255.100.42.133 | 225 | 255.165.144.185 | 205 | 253.108.42.152 | 185 | 9326297 | 3569795 | 2149713 | no | |
| emp005 | amit | m | Consultant | level 3 | 06-04-15 | 255.100.237.222 | 237 | 255.155.31.97 | 179 | 253.153.50.32 | 189 | 8207211 | 2156823 | 972960 | yes | |
| emp006 | kushal | m | developer | level 2 | 14-04-15 | 255.100.163.78 | 101 | 255.208.239.20 | 204 | 250.166.45.8 | 123 | 4449512 | 6016537 | 8218727 | yes | |
| emp007 | kasid | m | developer | level 2 | 15-07-15 | 255.100.98.30 | 216 | 255.192.230.141 | 210 | 250.200.180.50 | 140 | 8804780 | 9624421 | 1954318 | yes | |
| emp008 | shiv prakash | m | developer | level 2 | 27-07-15 | 255.100.14.25 | 219 | 255.207.72.65 | 89 | 254.120.200.227 | 192 | 1585871 | 3045069 | 5990606 | yes | |
| emp009 | vikram singh | m | developer | level 2 | 05-11-15 | 255.100.172.0 | 115 | 255.249.180.1 | 30 | 253.195.202.24 | 33 | 4084439 | 7681220 | 5259159 | no | |
| emp010 | sanjay | m | developer | level 2 | 30-11-15 | 255.100.253.89 | 41 | 255.185.117.185 | 112 | 250.238.22.85 | 110 | 1955243 | 8382193 | 2160678 | yes | |
| emp011 | abhi | m | developer | level 2 | 09-12-15 | 255.100.48.19 | 240 | 255.188.243.123 | 107 | 250.148.89.185 | 190 | 8669426 | 3335512 | 686697 | yes | |
| emp012 | ram dutt gupta | m | developer | level 2 | 13-01-16 | 255.100.6.237 | 226 | 255.243.82.191 | 70 | 251.181.141.117 | 38 | 2431177 | 8134807 | 4359854 | yes | |
| emp013 | khadak singh | m | developer | level 2 | 14-01-16 | 255.100.162.240 | 53 | 255.185.68.230 | 80 | 251.140.177.133 | 149 | 8233591 | 2579423 | 9326297 | no | |
| emp014 | gurmit singh | m | developer | level 2 | 21-01-16 | 255.100.151.104 | 91 | 255.113.152.181 | 244 | 251.158.91.194 | 208 | 2149713 | 8388766 | 8207211 | yes | |
| emp015 | chanderpal | m | team leader | level 3 | 22-01-16 | 255.100.180.62 | 116 | 255.184.58.100 | 187 | 251.177.75.246 | 199 | 972960 | 1543215 | 4449512 | no | |
| emp016 | aman | m | team leader | level 3 | 26-01-16 | 255.100.227.238 | 78 | 255.216.223.212 | 102 | 254.103.77.132 | 153 | 8218727 | 5170170 | 8804780 | yes | |
| emp017 | khursid | m | Consultant | level 4 | 13-04-16 | 255.100.17.55 | 76 | 255.219.45.193 | 117 | 254.230.197.103 | 64 | 1954318 | 1726693 | 1585871 | yes | |
| emp018 | rajeev | m | Consultant | level 3 | 30-05-16 | 255.100.201.12 | 100 | 255.174.233.163 | 55 | 254.178.17.201 | 72 | 5990606 | 2532979 | 4084439 | no | |
| emp019 | durgesh | m | developer | level 2 | 16-06-16 | 255.100.230.55 | 223 | 255.232.224.238 | 239 | 254.190.3.196 | 142 | 5259159 | 3219369 | 1955243 | yes | |
| emp020 | nahar singh | m | developer | level 2 | 23-08-16 | 255.100.195.126 | 87 | 255.232.54.151 | 37 | 252.204.185.196 | 168 | 2160678 | 5100890 | 9860231 | yes | |
| emp021 | ram kumar | m | developer | level 2 | 14-09-16 | 255.100.113.135 | 157 | 255.230.229.97 | 133 | 251.240.9.184 | 164 | 686697 | 2465757 | 1190065 | yes | |
| emp022 | sunder paal | m | developer | level 2 | 06-10-16 | 255.100.193.170 | 65 | 255.234.49.205 | 52 | 251.122.114.220 | 67 | 9860231 | 6135073 | 2233429 | no | |
| emp023 | maan singh aswal | m | developer | level 2 | 11-10-16 | 255.100.152.215 | 114 | 255.208.35.147 | 147 | 250.232.235.139 | 155 | 1190065 | 5807834 | 3569795 | no | |
| emp024 | rohit | m | developer | level 2 | 11-11-16 | 255.100.98.240 | 45 | 255.100.162.240 | 165 | 252.128.81.131 | 197 | 2233429 | 8669426 | 2156823 | no | |
| emp025 | rohit | m | developer | level 2 | 23-12-16 | 255.100.239.246 | 124 | 255.190.10.197 | 48 | 254.153.202.173 | 220 | 8669426 | 2431177 | 6016537 | no | |
| emp026 | sparsh | m | Project Manager | level 4 | 04-02-15 | 255.100.44.130 | 40 | 255.205.98.45 | 245 | 254.123.79.180 | 154 | 2156823 | 8233591 | 9624421 | yes | |
| emp027 | santosh | m | Project Manager | level 4 | 19-02-15 | 255.100.198.139 | 34 | 255.253.13.212 | 69 | 253.110.235.38 | 82 | 6016537 | 2149713 | 3045069 | no | |
| emp028 | santosh | m | Project coordinator | level 4 | 20-02-15 | 255.100.178.108 | 146 | 255.171.200.87 | 160 | 253.100.2.140 | 39 | 9624421 | 972960 | 7681220 | yes | |
| emp029 | punit khandelwal | m | Project coordinator | level 4 | 13-03-15 | 255.100.46.99 | 131 | 255.254.169.80 | 122 | 254.168.35.67 | 229 | 3045069 | 8218727 | 8382193 | yes | |
| emp030 | dinesh | m | Project coordinator | level 3 | 04-06-15 | 255.100.162.240 | 137 | 255.144.55.227 | 125 | 251.222.133.168 | 159 | 7681220 | 1954318 | 3335512 | yes | |
| emp031 | gulshan | m | Project coordinator | level 3 | 29-06-15 | 255.100.182.16 | 194 | 255.117.191.170 | 138 | 251.126.10.103 | 201 | 8382193 | 5990606 | 8134807 | yes | |
| emp032 | arvind kumar yadav | m | System Engineer | level 4 | 08-07-15 | 255.100.160.245 | 46 | 255.183.131.183 | 148 | 252.138.110.111 | 158 | 3335512 | 5259159 | 2579423 | no | |
| emp033 | nausad | m | System Engineer | level 3 | 10-07-15 | 255.100.117.52 | 59 | 255.200.187.240 | 120 | 250.180.210.153 | 163 | 8134807 | 2160678 | 8388766 | yes | |

Fig. 1: Dataset for intrusion detection system calculation

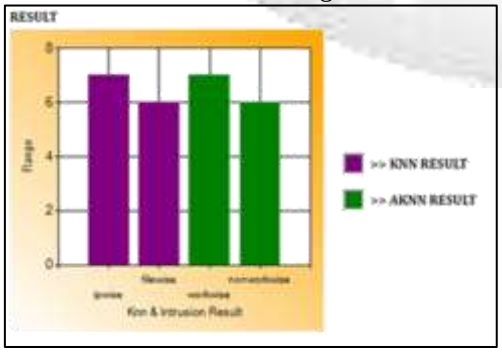


Fig. 2: knn & Intrusion result

## VI. CONCLUSION

In this paper we proposed advanced KNN based Dataset to identify the normal and attacked data. After that to enhance the algorithm we applied Genetic Algorithm which generates new Dataset using Mutation concept. We prove that, applying the combination of advanced KNN and Genetic Algorithm further improves the detection rate to a greater extent. Here, we present an Intrusion Detection System which detect and classify the attacks from normal users. We compute the false alarm rate in implementing this algorithm on an input dataset. Since a user profile does not

remain the same at all the time, we implement the Genetic Algorithm to observe the different variations possible in a user's profile. In this way we generate a new dataset, classify this new dataset using the advanced KNN and GP to compute the false alarm rate.

## REFERENCE

[1] Corinne Lawrence- "IPS – The Future of Intrusion Detection"- University of Auckland - 26th October 2004.
[2] Anita K. Jones and Robert S. Sielken –"Computer System Intrusion Detection A Survey "International Journal of Computer Theory and Engineering, Vol.2, No.6, December, 2010
[3] Vera Marinova-Boncheva-"A Short Survey of Intrusion Detection Systems using GP"-. Bulgarian academy of sciences.
[4] Carl Endorf, Eugene Schultz, Jim Mellander "Intrusion detection & prevention" by Written-published by McGraw-Hill.
[5] PeymanKabiri and Ali A.Ghorbani-"Research on Intrusion Detection and Response Survey"-International Journal of Network Security, Vol.1, No.2, PP.84–102, Sep. 2005
[6] Christopher Low –"Understanding Wireless attacks &detection "-GIAC Security Essentials Certification (GSEC) Practical Assignment 13 April 2005 -SANS Institute InfoSec Reading Room.
[7] "Global Information Assurance Certification Paper"-Copyright SANS Institute Copyright SANS Institute Author Retains Full Rights
[8] "SANS penetration testing copyright by SANS"-Copyright SANS Institute Author Retains Full Rights.
[9] Sriram Sundar Rajan, Vijaya Krishna Cherukuri-"An Overview of Intrusion Detection Systems".
[10] Asmaa Shaker Ashoor, Prof. Sharad Gore – "Importance of Intrusion Detection System"-International Journal of Scientific & Engineering Research, Volume 2, Issue 1, January-2011.
[11] Paul Innella- "The Evolution of Intrusion Detection Systems"-Tetrad Digital Integrity, LLC. International Journal of Security, Privacy and Trust Management (IJSPTM) Vol 4, No 1, February 2015 44
[12] "Intrusion Detection and Intrusion Prevention using various algorithm "-Ed Sale VP of Security Pivot Group, LLC.
[13] Shankar Sharan Tripathi, Sonu Agrawal- "A Survey on Enhanced Intrusion Detection System in Mobile Ad hoc Network"-International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 7, September 2012.