IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Analysis of different Hybrid methods for Intrusion Detection System

Durgesh Srivastava[1]*, Rajeshwar Singh[2], Vikram Singh[3]

[1]Department of CSE, IKG Punjab Technical University, Jalandhar, Punjab, India
[2], Department of ECE, DOABA Group of Colleges Rahon, SBS Nagar, Punjab, India
[3]Department of CSE Chaudhary Devi Lal University, Sirsa, Haryana, India.

Corresponding Author: dkumar.bit@gmail.com Tel.: +91-9996745993

DOI: https://doi.org/10.26438/ijcse/v7i5.757764 | Available online at: www.ijcseonline.org

Accepted: 23/May/2019, Published: 31/May/2019

Abstract- Critical incidents targeting National Critical Infrastructures are happening more and more often. Attacks, that happens to be both more sophisticated and persistent, can even replicate life. As per CERT-In's data, the number of cyber security incidents reported in the years: 2014-16 are more than 45000 and in 2017 (till June) are approx 27,482. Wannacry, Erebus & Petya are some big cyber-attacks, which crippled more than 10,000 organizations and 200,000 individuals in over 100 countries. From the above data, it's notable that the number of cyber security incidents has been growing steadily in India. The goal of this examination is to survey the relative performance of some notable hybrid classification techniques. We used KDD CUP 99 data to play out a controlled experiment in which the data characteristics are efficiently changed to present defects, for example, nonlinearity, multi-co-linearity, unequal covariance, and so forth. Our analyses recommend that datasets attributes significantly impact the classification execution of the strategies. Here we created and analyzed the diverse hybrid strategies in soft computing such as GWO-EBG, GWO-KNN, GWO-SVM and GWO-GRNN. The results of the diverse hybrid strategies can help in the structure of classification frameworks in which several classification techniques can be utilized to expand the reliability and consistency of the classification.

Keywords- Intrusion detection systems (IDS), SVM, Gray wolf optimizer (GWO), Entropy Based Graph, KNN etc.

I. INTRODUCTION

Classification issues is one of the main errand in data mining and AI, which target request everything in informational index into various collections subject to the information outlined by its properties. It is complicated to disconnect the features which are valuable, without past learning. From time to time, the datasets contain significant, insignificant and redundant attributes. So, the excess and unimportant properties can back off the classifier execution and they may even limit the classification accuracy because the search space becomes huge. Reduction of attribute could deal with this issue by picking just pertinent attribute for classification [1]. The decrease set will improving the classifier execution and giving a quicker and more cost effective order, which prompts acquire practically identical or even best classification accuracy from using all properties. GWO is one of the as of late proposed swarm knowledge based calculations, which is created by Mirjalili et al. [2] in 2014. An Intelligent GWO method is motivated by grey wolves scanning for the ideal route for chasing preys. GWO algorithm applies the same mechanism in nature, where it follows the pack hierarchy for sorting out the distinctive jobs in the wolves pack. In GWO, pack's individuals are separated into four gatherings dependent on the kind of the wolf's job that help in advancing the hunting process explain in section 2.

In this paper, KDD CUP 1999 data-sets [3][4] are used to experiment for Intrusion detection System (IDS) [5] [6] and the comparison of performances of the different classification techniques using GWO like Entropy based graph (GWO-EBG), K-nearest neighbor (GWO-KNN) [7], Generalized regression neural network (GWO-GRNN) [8] [9] and Support vector machine (GWO-SVM) [10] regarding performance measure like accuracy, sensitivity, specificity, PPV, NPV, FPR, FNR, FDR, F-measure and MCC [11][1812].

This paper is sorted out as pursues; some fundamental ideas of Gray wolf optimization (GWO) in segment 2, Different classifier techniques like entropy based graph, k nearest neighbor, support vector machine and generalized regression neural network are explore in section 3, explanation of the results and discussion in section 4. Then conclusion of analysis of intrusion detection system using different hybrid methods is in section 5.

II. GRAY WOLF OPTIMIZER (GWO)

Grey wolf optimization is a swarm intelligent method which mimics the leadership progression of wolves is familiar for their group hunting. Grey wolf for the most part want to live in a pack and they have a firm social predominant pecking order; the leader is a male or female, known as Alpha (α). The alpha is for the most part in charge for making

decisions. The requests of the prevailing wolf ought to be trailed by the group. The Betas (β) are foot soldier wolves which assist the alpha (α) in fundamental administration. The beta is a counselor to alpha and discipliner for the pack. On the off chance that a wolf is neither an α (or) β , is called δ . Delta (δ) wolves command omega and reports to alpha and beta [13] [14] [15].

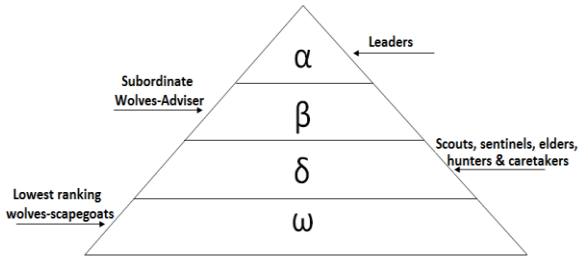

Fig. 1: Dominant pecking order of grey

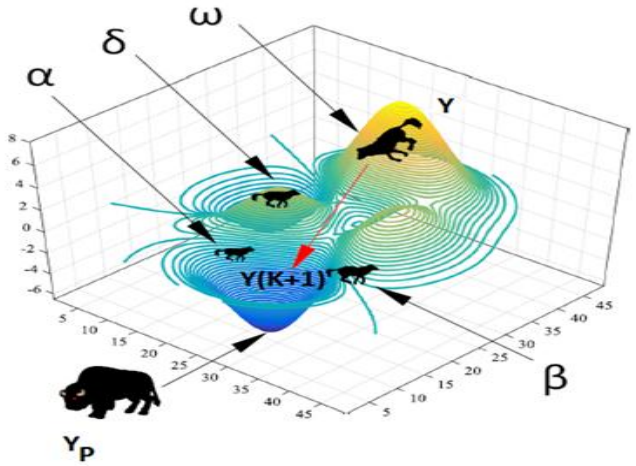

Fig. 2: Alpha (α), beta (β), delta (δ), omega (ω) wolf (top to bottom) and prey (Yp) are characterized in GWO

In GWO mathematical representation, the fitness solution is known as the alpha (α). The second, third and fourth most excellent solutions are named β, δ & ω individually. Here, the hunting is guided by α, β, δ and ω. The concepts of α, β, δ and ω are illustrated in figure 1. Note that the goal is to locate the base in this hunt scene. It might be found in this assume α is the wardrobe answer for the base, trailed by β and δ. Whatever remains of arrangements are considered as ω wolves. There is only one ω in figure 2, yet there can be more. The hunting method and the social pecking order of wolves are mathematically demonstrated keeping in mind the end goal to create GWO and perform optimization.

III. DIFFERENT CLASSIFICATION METHODS

A. Entropy Based Graph (EBG)

In this section, entropy based graph is utilized for the effective classification of data into normal or intrusion. Entropy values are estimated for the chosen features set and after that mean entropy is ascertained from the entropy esteems. At that point the mean entropy is kept as a threshold value for the valuable classification of data into normal or abnormal. The sum of every data makes a random variable for which expected esteem or average is the entropy. Entropy is evaluated for the informational collections by utilizing condition (1).

$$E_y(set) = \sum_{i=1}^{M} P(val_i)\ \log_2\{P(val_i)\} \quad (1)$$

Where, $P(val_i)$ is the probability of chosen of i^{th} feature. Mean esteem is ascertained for all the entropy esteems after the calculation of entropy measure for the features. In the proposed classification this entropy measure is taken as a threshold value for the valuable classification of data into normal or intrusion. The mean entropy measure is computed by the condition (2).

$$M(E_y(set)) = \frac{E_y(set)}{M} \quad (2)$$

Where, M is the total number of entropy measures. The entropy based graph generation is depends on the mean entropy measure. The algorithm of proposed entropy based graph classification is specified in figure 3.

In entropy based graph generation for the categorization of data into normal or intrusion, the input is $Y = < y_1, y_2, y_3, ... y_k >$ is the processed data and it is taken as an input. At first for every data $y_k \in Y$ ascertains the entropy utilizing condition (1) and mean entropy esteem is then computed by utilizing the condition (2). The mean entropy value is taken as a threshold value for the classification of data. The example diagram for classification of data into normal or intrusion using entropy based graph classification is given in figure 3[16], [17].

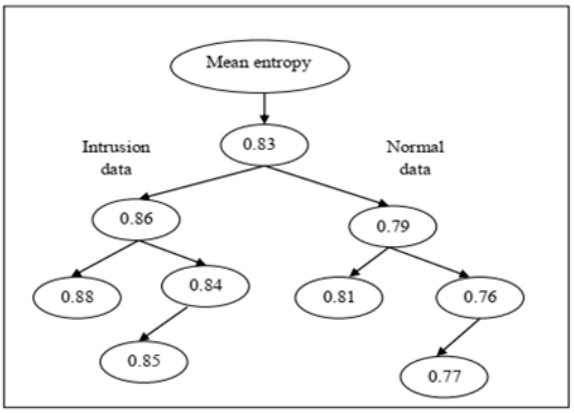

Fig. 3: Example diagram of entropy based graph classification

In the classification, the entropy of chosen feature esteem is higher than the threshold value then the data is located in the left side of the graph i.e.) intrusion data and if the entropy of selected feature value is lesser than the threshold value is placed in the right side correspondingly i.e.) normal data.

B. Support Vector Machine (SVM)

SVMs are set of related supervised learning strategies utilized for classification and regression [18]. They belong to a set of generalized linear classification. A special property of SVM is, SVM minimize the observational classification mistake and maximize the geometric edge simultaneously. So SVM knew as maximum margin Classifiers and SVM depends on the Structural risk Minimization (SRM). SVM map input vector to a higher dimensional space where a maximal isolating hyperplane is developed. Two parallel hyperplanes are developed on each side of the hyperplane that separate the data. The isolating hyperplane is the hyperplane that amplify the separation between the two parallel hyperplanes. A supposition that is made that the bigger the edge or separation between these parallel hyperplanes the better the speculation error of the classifier will be [1] [18]. Figure 4 delineate the SVM procedure.

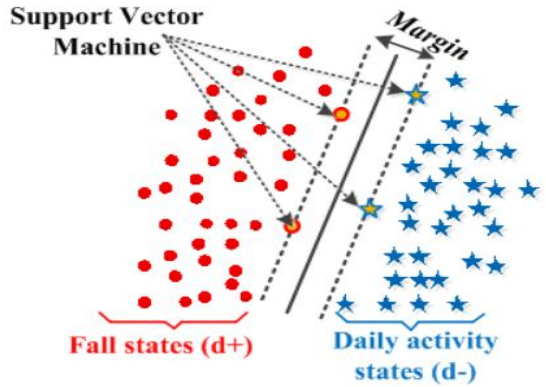

Fig. 4: SVM process

C. K-Nearest Neighbor (KNN)

KNN algorithms have been used since 1970 in many applications like statistical estimation and pattern recognition etc. It is a basic algorithm, which stores all cases and arranges new cases dependent on similitude measure. KNN algorithm also called as 1) case based thinking 2) k nearest neighbor 3) model based reasoning 4) occurrence based learning 5) memory based reasoning 6) lazy learning.. It is a non parametric classification strategy which is two types1) structure less NN methods 2) structure based NN procedures. In structure less NN strategies, entire information is ordered into training and test sample data. From training point to test point distance is calculated, and the point with lowest distance is called nearest neighbor. Structure based NN systems are based with respect to structures of data like orthogonal structure tree (OST), ball tree, k-d tree, pivot tree, closest future line and focal line [7]. Nearest neighbor classification is utilized mostly when every one of the attributes are continues. K nearest neighbor algorithm is -

Step: 1 Find the K training instances which are nearest to obscure instance.
Step: 2 choose the most ordinarily happening classification for these K instance.

There are different methods for estimating the similarity between two cases with n attribute values. Each measure has the accompanying three requirements. Let dist (A, B) be the separation between two points A, B then,

1) dist (A, B) >= and dist(A, B)-0 iff A-B
2) dist (A, B) = dist(B, A)
3) dist (A, C) <= dist(A, B) + dist(B, C)

Property 3 is called as "Triangle in equality". It expresses that the most limited distance between any two point is a straight line. Most basic separation estimates utilized is Euclidean distance. For constant factors Z score standardization and min max normalization are used [17].

D. Generalized Regression Neural Network (GRNN)

GRNN, as proposed by Donald F. Specht in [18] falls into the category of probabilistic neural networks. It is a memory-based network that provides estimates of continuous variables and converges to the underlying (linear or nonlinear) regression surface. This general regression neural network (GRNN) is a one-pass learning algorithm with a highly parallel structure. Even with sparse data in a multidimensional measurement space, the algorithm provides smooth transitions from one observed value to another. The algorithmic form can be used for any regression problem in which an assumption of linearity is not justified. The parallel network form should find use in applications such as learning the dynamics of a plant model for prediction or control [18]. Figure 5 display the structure of generalized regression neural network.

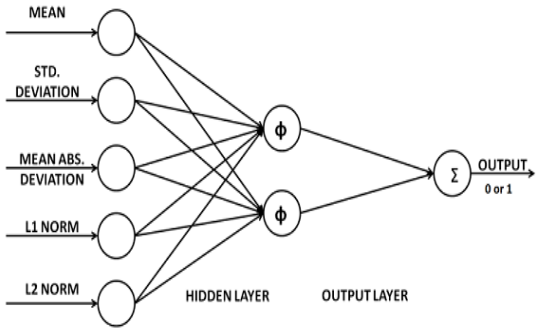

Fig. 5: Generalized Regression Neural Network Structure

IV. RESULT AND DISCUSSION

Different hybrid methods for Intrusion Detection system were implemented in the working platform of MATLAB R2014a with machine configuration as takes after: Processor: Intel core i5, CPU Speed: 3.20 GHz, OS: Windows 7 and RAM: 4GB. In this section, the experimental results accomplished with different hybrid methods. The publicly available KDD-CUP 99 dataset [2] [19] was utilized to assess the classification of data into normal or intrusion using different hybrid classification methods. The performance of different classification methods with Gray Wolf Optimization (GWO) like Entropy based graph (GWO-EBG), K-nearest neighbor (GWO-KNN), Support vector machine (GWO-SVM) and Generalized regression neural network (GWO-GRNN) are compared with different measure in percentage like accuracy, sensitivity, specificity, PPV, NPV, FPR, FNR, FDR, F-measure and MCC [11] [12]. The comparison results regarding of various performance measures are depicted in table 1, using different number of testing data sets. The comparison graph of performances measures are also displayed in figures 6-15.

Table 1: Performance table of different Hybrid methods with different numbers of testing data

Size of Input Data (Testing)	MEASURES	GWO-EBG	GWO-KNN	GWO-SVM	GWO-GRNN
1000	Accuracy (%)	95.7	77.50	79.00	77.80
	Sensitivity (%)	92.17	68.52	69.84	69.17
	Specificity (%)	99.03	91.73	93..33	90.75
	PPV (%)	98.89	92.92	94.25	91.81
	NPV (%)	93.07	64.78	66.42	66.24
	FPR (%)	0.97	8.27	6.67	9.25
	FNR (%)	7.84	31.49	30.16	30.83
	FDR (%)	1.11	7.08	5.75	8.19
	F measure (%)	95.41	78.87	80.23	78.90
	MCC (%)	91.58	58.96	61.91	59.00
2000	Accuracy (%)	94.80	77.05	79.30	77.20
	Sensitivity (%)	90.07	67.31	69.47	67.71
	Specificity (%)	99.32	92.80	94.52	91.95
	PPV (%)	99.21	93.80	95.15	92.90
	NPV (%)	91.28	63.70	66.67	64.70
	FPR (%)	0.68	7.20	5.48	8.05
	FNR (%)	9.93	32.69	30.54	32.30
	FDR (%)	0.79	6.20	4.85	7.10
	F measure (%)	94.42	78.38	80.31	78.33
	MCC (%)	89.94	58.80	62.89	58.62
3000	Accuracy (%)	94.93	76.80	78.60	76.57
	Sensitivity (%)	90.14	67.01	68.63	67.07
	Specificity (%)	99.48	92.46	94.19	91.19
	PPV (%)	99.40	93.43	94.86	92.15
	NPV (%)	91.40	63.66	65.75	64.26
	FPR (%)	0.52	7.54	5.81	8.81
	FNR (%)	9.86	33.00	31.37	32.93
	FDR (%)	0.60	6.57	5.14	7.86
	F measure (%)	94.54	78.04	79.65	77.63
	MCC (%)	90.20	58.27	61.71	57.33
	Accuracy (%)	94.80	75.95	76.83	76.03
	Sensitivity (%)	89.44	65.30	65.84	65.63

4000	Specificity (%)	99.71	93.03	94.96	91.96
	PPV (%)	99.65	93.77	95.57	92.60
	NPV (%)	91.15	62.57	62.74	63.57
	FPR (%)	0.29	6.97	5.04	8.04
	FNR (%)	10.57	34.70	34.16	34.37
	FDR (%)	0.35	6.24	4.43	7.40
	F measure (%)	94.27	77.00	77.97	76.82
	MCC (%)	89.97	57.32	59.54	56.88
5000	Accuracy (%)	94.60	75.62	76.66	75.74
	Sensitivity (%)	89.12	64.91	65.71	65.27
	Specificity (%)	99.65	93.27	94.84	92.23
	PPV (%)	99.58	94.09	95.48	92.97
	NPV (%)	90.85	61.73	62.50	62.78
	FPR (%)	0.35	6.73	5.16	7.78
	FNR (%)	10.88	35.09	34.29	34.73
	FDR (%)	0.42	5.92	4.52	7.03
	F measure (%)	94.06	76.82	77.84	76.70
	MCC (%)	89.60	56.98	59.25	56.61

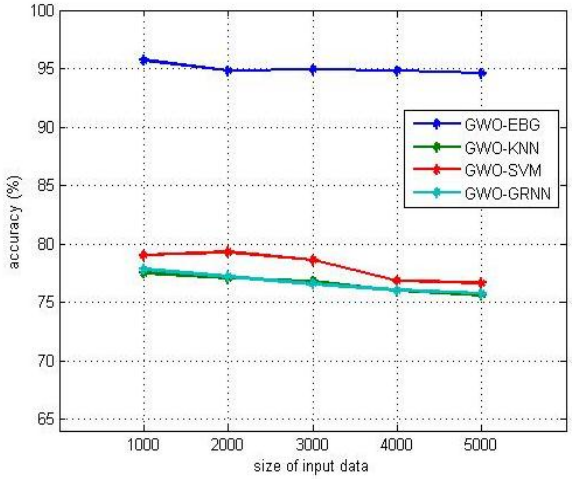

Fig. 6: Comparison Graph in Terms of Accuracy

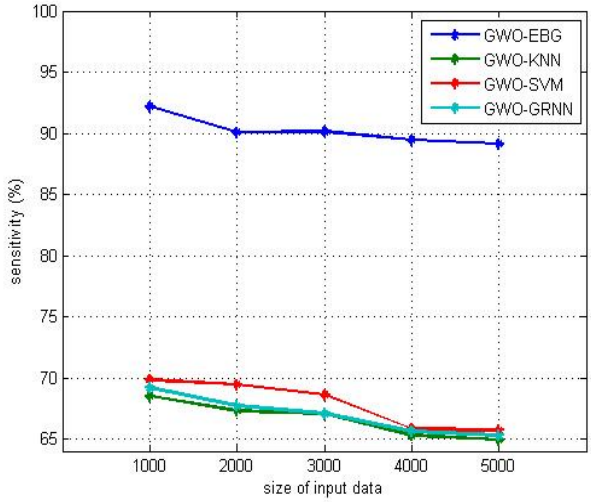

Fig. 7: Comparison Graph in Terms of Sensitivity

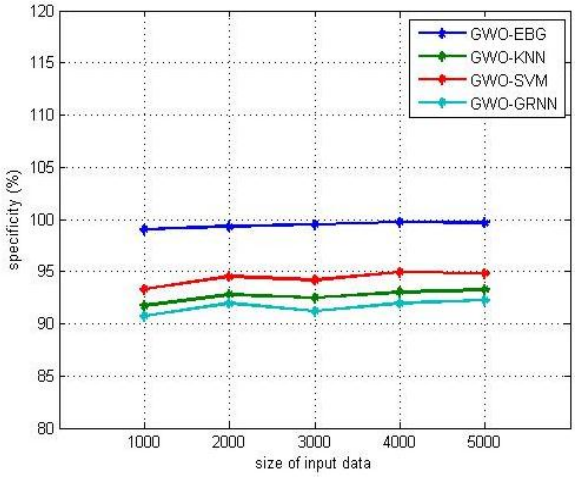

Fig. 8: Comparison Graph in Terms of Specificity

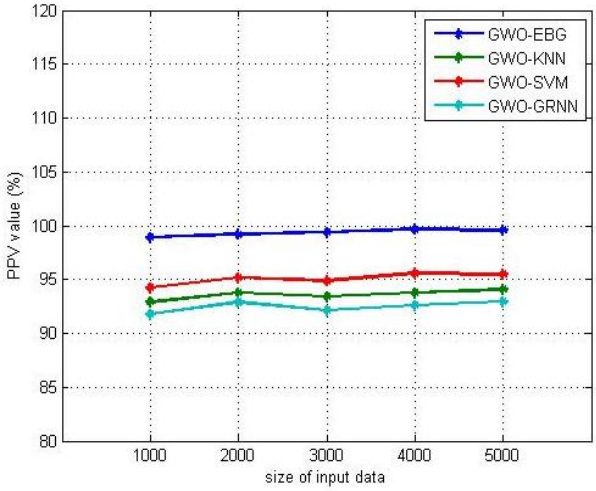

Fig. 9: Comparison Graph in Terms of PPV value

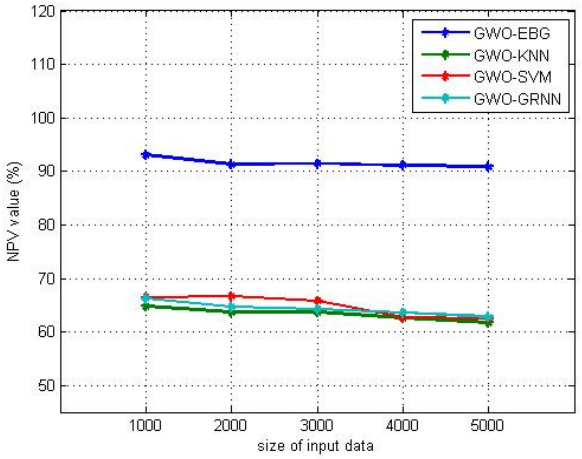

Fig. 10: Comparison Graph in Terms of NPV Value

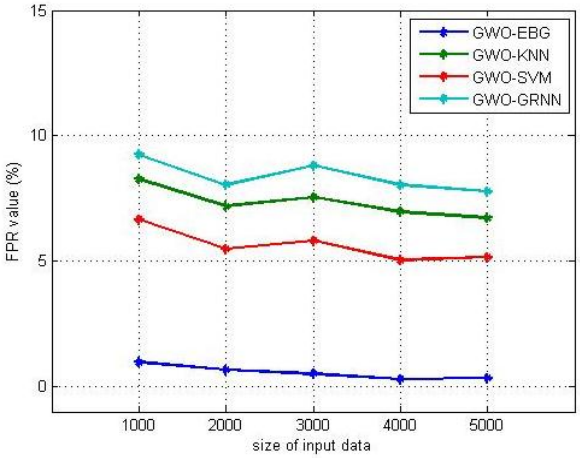

Fig. 11: Comparison Graph in Terms of FPR Value

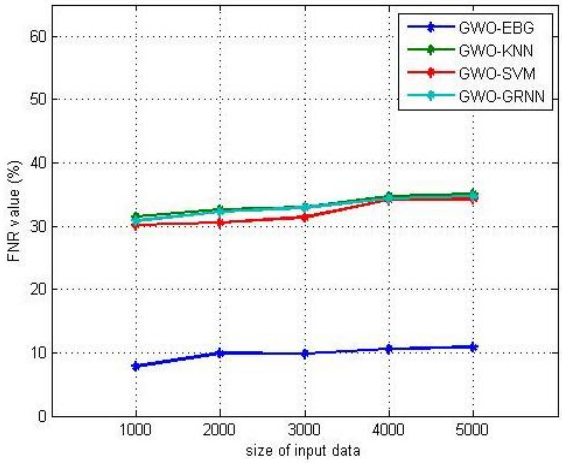

Fig. 12: Comparison Graph in Terms of FNR Value

The statistical metrics of sensitivity, specificity, accuracy and others can be expressed in the terms of TP, FP, FN and TN esteems. The performance of different hybrid methods

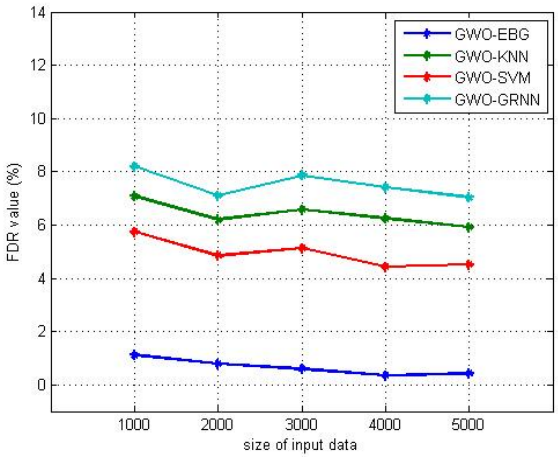

Fig. 13: Comparison Graph in Terms of FDR value

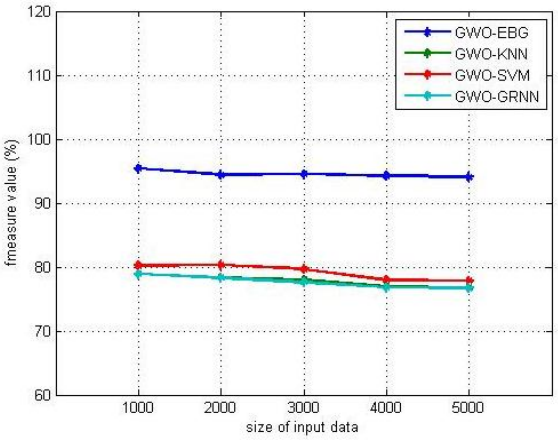

Fig. 14: Comparison Graph in Terms of F-measure Value

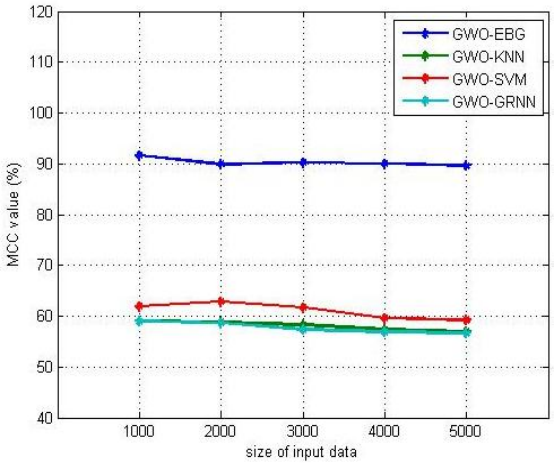

Fig. 15: Comparison Graph in Terms of MCC value

are studied by utilizing the statistical measures mentioned in this section. The figure 6 illustrates the GWO and entropy based graph (GWO-EBG) hybrid classification have an accuracy slightly higher compared to other hybrid

methods. The figure 7 and 8 illustrates the GWO-EBG classification have higher sensitivity and specificity compared to other hybrid methods. Figure 9, 10, 14 and 15 illustrates the GWO and entropy based graph (GWO-EBG) hybrid method, which have slightly higher Positive Predictive Value (PPV), Negative Predictive Value (NPV), F measure value (FMV) and Mathew's correlation coefficient (MCC) value than other hybrid methods. Figure11, 12 and 13 represents that the GWO-EBG hybrid method has lower false positive rate (FPR), false negative rate (FNR) and false discovery date (FDR) compared to other hybrid methods like GWO-KNN, GWO-SVM and GWO-GRNN.

V. CONCLUSION

In this paper, the significance of security to clients on the system either personal clients or in associations has been underlined on different occasions; the gravity of this significance keeps on requiring repeat of newer and updated researches in the zone. So, it is discussed different classification method with an intelligent Gray Wolf Optimization (GWO) technique for network intrusion, which is a nature inspired technique used for key feature selection. The KDD CUP 99 data sets are pre-processed, and the features are optimally chosen by using optimize GWO algorithm. So, data sets are reduced from 41 features to 24 features. After getting optimal feature sets, applied different classification techniques to make it hybrid like Entropy based graph (EBG), K-nearest neighbor (KNN), Support vector machine (SVM) and generalized regression neural network (GRNN) to classify the data into normal and intrusion classes. Table 1 and figures 6-15 shows comparison results of different hybrid methods like GWO-EBG, GWO-KNN, GWO-SVM and GWO-GRNN by using different numbers of data sets with respect to different measures like accuracy, sensitivity, specificity, PPV, NPV, FPR, FNR, FDR, F-measure and MCC value.

REFERENCES

[1] D.K. Srivastava, K. S. Patnaik and L Bhambhu, "*Data Classification: A Rough - SVM Approach*", in Contemporary Engineering Sciences, Vol. **3** no. **2, 2010**, pp 77 – 86.

[2] S. Mirjalili, S. M. Mirjalili, A. Lewis, "*Grey wolf optimizer"*, Advances in Engineering Software, vol. **69, 2014**, pp. 46-61.

[3] KDD cup 1999 data, http:// kdd.ics.uci.edu/ databases/kddcup99/kddcup99.html .

[4] Durgesh Srivastava, Nachiket Sainis and Dr. Rajeshwar Singh, "*Classification of various Dataset for Intrusion Detection System*", in International Journal of Emerging Technology and Advanced Engineering, Volume **8**, Issue **1**, January **2018**.

[5] H. Günes, Kayacık, A, NurZincir-Heywood, Malcolm I. Heywood, "*Selecting Features for Intrusion Detection: A Feature Relevance Analysis on KDD 99 Intrusion Detection Datasets*", in Third Annual Conference on Privacy, Security and Trust, October 12-14, **2005**

[6] Chet Langin, Shahram Rahimi, "*Soft computing in intrusion detection: the state of the art*", J ambient Intel Human Compute (**2010**), 1:133–145, Springer.

[7] Lin, Wei-Chao, Shih-Wen Ke, and Chih-Fong Tsai. "*CANN: An intrusion detection system based on combining cluster centers and nearest neighbors*", in Knowledge-based systems, **2015**.

[8] Donald F. Specht, "*A General Regression Neural Network*", in IEEE transactions on neural networks, Vol. **2**, No. **6**. November **1991**.

[9] Benmessahel, Ilyas, Kun Xie, and Mouna Chellal. "*A new evolutionary neural networks based on intrusion detection systems using multiverse optimization*", Applied Intelligence **2017**.

[10] Durgesh Srivastava, L Bhambhu, "*Data classification using support vector machine*" Journal of Theoretical and Applied Information Technology, **12(1), 2010**.

[11] Alaa Tharwat, "Classification assessment methods", in Applied Computing and informatics, **2018.**

[12] Okeh UM and Okoro CN, "*Evaluating Measures of Indicators of Diagnostic Test Performance: Fundamental Meanings and Formulars*", in Journal of Biometrics & Biostatistics, Vol.**3**, Issue **1**, **2012**

[13] Hossam Faris, Ibrahim Aljarah, "*Grey wolf optimizer: a review of recent variants and applications*", in Neural Computing and Applications, **2018**.

[14] Jitendra Kumar, Satish Chandra, "*Intrusion detection based on key feature selection using Binary GWO*", in International conference on computing for sustainable global development, **2016**.

[15] Qiang Li, Huiling Chen, Hui Huang, "*An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis*" , in Computational and Mathematical Methods in Medicine, Article ID 9512741, 15 pages, **2017**.

[16] Durgesh Srivastava, Rajeshwar Singh and Vikram Singh, "*Performance Evaluation of Entropy Based Graph Network Intrusion Detection System (E-Ids)*", in Jour of Adv Research in Dynamical

& Control Systems, Vol.- **11, 02**-Special Issue, **2019**

[17] Durgesh Srivastava, Rajeshwar Singh, Vikram Singh, "*An Intelligent Gray Wolf Optimizer: A Nature Inspired Technique in Intrusion Detection System (IDS)*", in Journal of Advancements in Robotics. **2019**; **6(1):** 18–24p

[18] Durgesh Srivastava, L Bhambhu, "*Data classification using support vector machine*" Journal of Theoretical and Applied Information Technology, **12(1), 2010**.

[19] Ebrahim Bagheri, Wei Lu, Mahbod Tavallaee and Ali A. Ghorbani , "*A Detailed Analysis of the KDD CUP 99 Data Set*", in IEEE Symposium on Computational Intelligence for Security and Defense Applications, **2009**.

Biography:

Mr. Durgesh Kumar Sriavastava is a PhD research scholar of Department of CSE, IKG Punjab Technical University, Jalandhar, Punjab, India. He has received B.Tech degree in Information & Technology (IT) from MIET, Meerut, UP, INDIA in 2006 and ME in software engineering from Birla Institute of Technology (BIT), Mesra, Ranchi, Jharkhand, INDIA) in 2008. Currently, he is an Assistant Professor (AP) at BRCM CET, Bahal, Bhiwani, Haryana, INDIA. His interests are in Machine Learning, Soft Computing, Pattern recognition, Software engineering, Modeling and design etc.

Dr. Rajeshwar Singh working as a Professor and Director in DOABA Group of Colleges, Punjab, India. He has received his PhD in electronic & communication engineering from Vinoba Bhave University, Bihar, India and M.Tech in Digital System from University of Allahabad, UP. His research area are Soft computing, Machine Learning, Communication, Wireless Sensor Network etc.

Dr. Vikram Singh, working as a Professor in the Department of Computer Sci., Chaudhary Devi Lal University, Sirsa, Haryana, India. He has received MCA, PhD in Computer Science from K.U. Kurukshetra, Haryana, India. Also hold the additional charges of Dean of Colleges and Director of University Computer Centre at the university. His research interest areas are Data Mining, Wireless Sensor Network, Simulation and Modeling, Software Development etc.

