Accurate Intrusion Detection Based On Feature Optimization Using Plant Grow Algorithm

Moti Kumari, Monika Raghuvanshi
Department of Computer Science and Engineering
Bhabha Engineering Research Institute, MP, Bhopal, India
motikumari583@gmail.com, monipriya@gmail.com

Abstract: *The process of features reduction enhanced the performance of the intrusion detection system. Nowadays used various features reduction algorithms are used for static as well as dynamic features reduction. The feature reduction technique behaves in dual mode. The reduction of features cannot have fixed how many features are reducing for the better detection process of intrusion. The process of features reduction used plant grow optimization algorithm and classification using support vector machine algorithm.*

Keywords: -IDS, Feature Matrix, SVM, Accuracy, Precision, Recall, KDDCUP99, Machine Learning, features, Detection.

I. Introduction

An interruption recognition framework progressively screens the occasions occurring in a checked framework and chooses whether these occasions are symptomatic of an assault or constitute an honest to goodness utilization of the framework. Figure delineates the association of an IDS where strong bolts show information control stream while specked bolts demonstrate a reaction to meddlesome exercises [1-3]. As a rule, IDSs fall into two classes as their recognition strategies, particularly (i) abuse identification and (ii) irregularity discovery. Abuse location recognizes interruptions by coordinating watched information with pre-characterized depictions of nosy conduct. In this way, understood interruptions could be recognized proficiently with a low false-positive rate.

Consequently, the approach is generally embraced in the larger part of business frameworks. Nonetheless, interruptions are typically polymorph and develop constantly. Abuse location flop effortlessly when confronting obscure interruptions [7-9]. One approach to delivering this issue is routinely redesigning the information base, either physically tedious and relentless or naturally assisting directed learning calculations. Lamentably, datasets for this pure-posture are typically costly to get ready, as they require the naming of every case in the dataset as ordinary or a kind of interruption. Another approach to delivering this issue is to take after the irregularity discovery show talked about by Denning [12-14]. In the rest of the research, plant grow optimization is described in section II, proposed algorithm explained in section III, Experimental result Analysis discussed in section IV and finally, conclusion and future work discussed in section V.

II. Plant Grow Optimization

The PGO takes the problem's solution space as the growth area of the artificial plant, in which one point of the plant is one potential solution to the problem. The algorithm searches the optimal point in the solution space through two behaviours [5-6]:

1. Producing new points by branching to search the optimal area where the optimum solution is;
2. Growing leaves around the branch point to find the accurate solution in the local area;

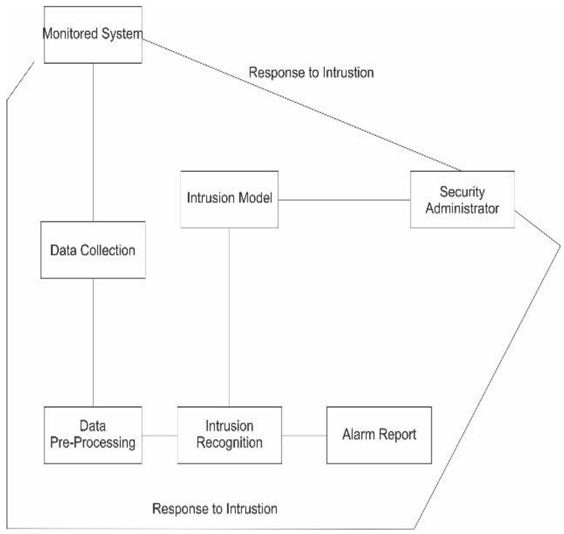

Figure 1: Organization of a generalized intrusion detection system [4].

Given the definitions of the preceding section, formally the Plant Growth Optimization is[10-11]:

Start

Initialize:

Set NG=0 {NG is the generations counter}

Set NC=0 {NC is the convergence counter}

Set NM=0 {NM is the Mature points counter}

Set the upper limit of the branch points N and initialize other parameters.

Select N0 branch points at random and perform leaf growth.

Assign morphogen

Calculate the eligibility of the leaf point.

Assign the concentration of the morphogen of each branch point.

Branching

Select two critical values between 0 and 1 randomly and dispose of them.

Produce new points by branching in four modes.

Selection mechanism

Perform leaf growth in all the points.

Pick out the mature branch points, the number of which is $k\,(0 \leq k \leq N)$, by the maturity mechanism.

Set NM = NM + k

Produce a new point in the centre of the crowded area and select the best point to substitute the crowded points.

Eliminate the lower competition ability branch points and select N branch points for the next generation.

Competition

Compare the current points with the mature points and get the best fitness value of f_{max}

Set: $NG = NG + 1$

$If\ \left(f_{max} < f_{\max_{old}}\right) Set: f_{max} = f_{\max\_old}$

If $\left(\left|f_{max} - f_{\max_{old}}\right| < \varepsilon\right) Set: NC = NC + 1$

else

Set: NC = 0

else

Set: NC = NC+1

Check the termination criteria:

$If\ (NG < NG_{max} \&\& NC < NC_{max} \&\& NM < NM_{max})$

Go to step 2

else

Exit

Stop

One execution of the procedure from step2 through step6 is called a generation or a cycle.

III. Proposed Algorithm

Feature reduction and classification of intrusion data is a major issue. For the reduction of features used various optimization techniques. This article used the plant grow optimization technique for the reduction and selection of features. The plant grows the process of development of plants inspires optimization algorithm. The development of plants is divided into three sections as described below.

1. Morphogen
   In the case of morphogen, check the status of plants for growing.
2. Branching
   In the case of branching, check the section condition of the new leaf policy
3. Termination
   Termination is the final process of plant theory. The termination process gives the optimal solution to the given problem

The following parameter is used for the path process, x1, x2,…………..,xn is the path component of the robot. W is the Wight factor for the path, T is the value of morphogen, c1 and c2 is the contour value of the path.

Step1. Define the value of path set S1{x1,x2,…………xn} with population

Assign the value of contour and weight of path C1=0, C2=0 and W=0.

a. Morphogen selection of plant function

$$F(s) = \frac{(Ffd - Fpf)}{Fd * fp}, wi\epsilon S(x1, x2 \ldots\ldots xn) \ldots\ldots\ldots\ldots\ldots (1)$$

Here Ffd is the process features set, and Fpf is the final features set of plant, and w is the set of the path component of sum sets

The features set the value of the branch $F = \{fa1 \ldots\ldots an\}$. These branch values proceed for the estimation competition condition of the local leaf.

$$Fcom = \begin{cases} \dfrac{(T_i)^{\alpha}\left(LI_i^{S_j}\right)^{\beta}}{\sum_{g \notin S_j}(\tau_g)^{\alpha}\left(LI_g^{S_j}\right)^{\beta}} if\ \ i \notin S_j \\ 0 \qquad\qquad otherwise \end{cases} \quad (2)$$

Here T is a target value of features, and LI is the value of features difference.

Step2. Branching condition

Input the selected path for the Competition

1. Calculate the value of relative features of C1 and C2
   $\text{Rf} = \frac{\text{LSI}}{\text{Wd}}$ Here Lsi the difference of intrusion features
2. The PGO estimate the optimal features for selection.
   $$FS = \begin{cases} \dfrac{\max\ (RF)\ -F(s)}{\max_{h=1:}(WS)} & if\ \ s_i \in f_j \\ 0 & otherwise \end{cases} \quad (3)$$
3. create the relative FS difference value of features set
   $$\text{Rd} = \sum_{\text{fd}=1}^{\text{n}} \sum_{\text{pf}=1}^{\text{m}} (\text{xi} - \text{fs}) \ldots\ldots\ldots\ldots\ldots\ldots (4)$$
4. if the value of Rd is zero, the features optimization process is terminated

step 3 Termination

. Where $Rd$is the relative difference of $T(i)$; $f_z$is the fitness value; standard deviation $S_z$and local density $D_z$are defined in formula (5):

$$\begin{cases} R_d = \sqrt{\dfrac{\sum_{i=1}^{n}(z(i) - E(z))^2}{(n-1)}} \\ f_z = \sum_{i=1}^{n}\sum_{j=1}^{n}(R - r(i,j))\, u(R - r(i,j)) \end{cases} \quad (5)$$

.

Defining $d(z(k), z(h))$as the absolute distance between the two-optimal path

$$d(z(k), z(h)) = \sqrt{(z(k) - z(h))(z(k) - z(h))} = \sqrt{(z(k) - z(h))^2}$$

$k = 1, 2, \ldots\ldots\ldots, N; h = 1, 2, \ldots\ldots\ldots, N$ and finally, the path is terminated.

step 4 Input of classifier (SVM)

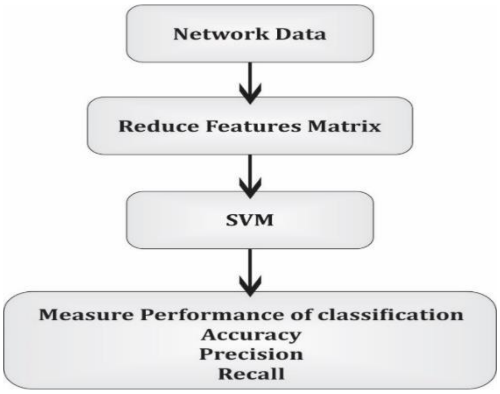

Figure 2: process block diagram of optimized features classification using support vector machine.

IV. Experimental and Result Analysis

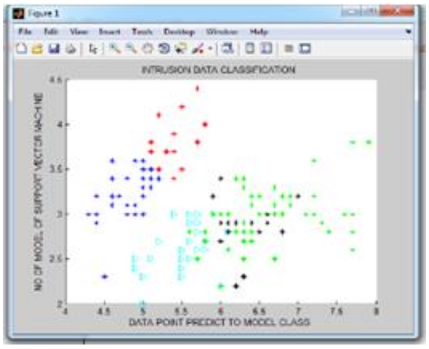

Figure 3: window shows that the number of attributes reduces value is 3, using the SVM method to enhance the performance of Intrusion Detection System Based on Feature Reduction using Plant Grow Optimization Algorithms.

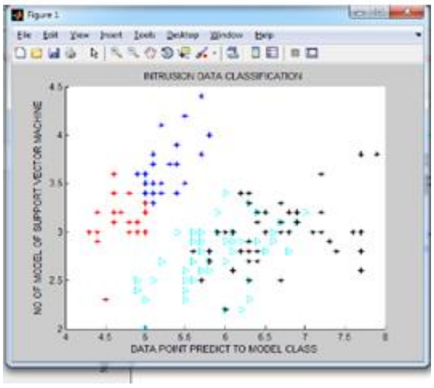

Figure 4: window shows that the number of attributes reduction values is 11, using the PROPOSED method to enhance the performance of Intrusion Detection System Based on Feature Reduction using Plant Grow Optimization Algorithms.

Table 1: Comparative output value of our implementation using SVM and Proposed Method with input number of attribute reduces 7, 11, and 18.

Method	Accuracy	Precision	Recall	Attribute
SVM	80.2659	78.2659	77.2659	7
PROPOSED	87.8709	82.8709	83.8709	
SVM	82.8709	80.8709	79.8709	11
PROPOSED	98.9401	97.9401	99.9401	
SVM	85.2659	80.2659	81.2659	18
PROPOSED	99.9401	98.9401	97.9401	

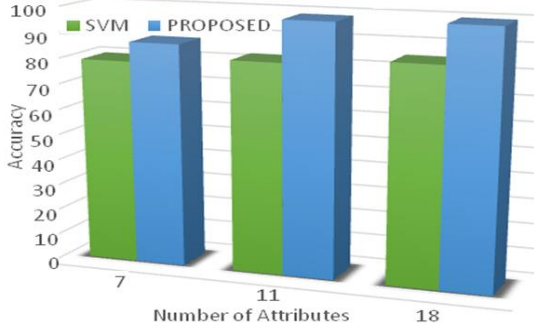

Figure 5: Comparative result graph based on the output value of Accuracyusing SVM and Proposed Method for input number of attribute reduces 7, 11 and 18 in Enhanced the performance of Intrusion detection System.

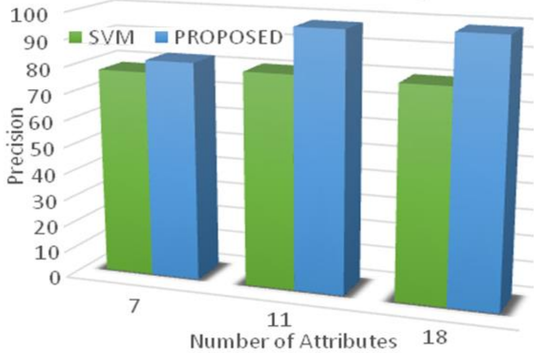

Figure 6: Comparative result graph based on the output value of Precision using SVM and Proposed Method for input number of attributes reduces 7, 11 and 18 in Enhanced the performance of Intrusion detection System.

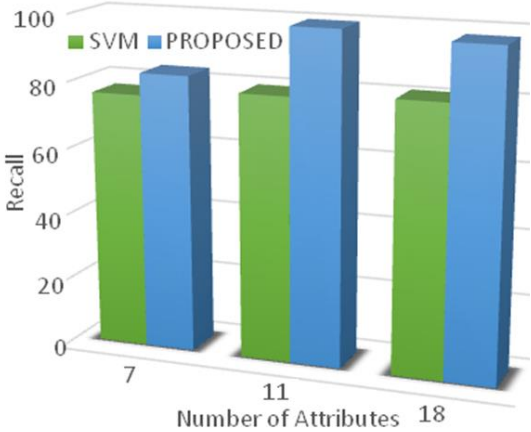

Figure 7: Comparative result graph based on the output value of Recall using SVM and Proposed Method for input number of attribute reduces 7, 11 and 18 in Enhanced the performance of Intrusion detection System.

V. Conclusions

The processing of network data is very complex and now required network features optimization. This dissertation used the plant grow optimization technique for the reduction of features. The plant grows the behaviour of plant kingdom inspires optimization technique algorithm. The reduced attribute classified by well know classifier is called a support vector machine. The combination of support vector machine

ISSN: 2319-7900 www.ijact.org Volume-X, Issue-IV, August 2021

and plant grow optimization performs very well in compression of the previous feature reduction technique. The plant growth optimization with a support vector machine is better than the feature reduction and classification SVM process. The proposed algorithm is very efficient for dynamic attributes for the classification problem. The detection and classification process is better than the previous method. In future, uses a multi-agent glowworm optimization algorithm.

References

[1]. Preeti Singh and Amrish Tiwari "An Efficient Approach for Intrusion Detection in Reduced Features of KDD99 using ID3 and classification with KNNGA", IEEE, 2015, Pp 445-452.

[2]. Gaby Abou Haidar and Charbel Boustany "High Perception Intrusion Detection Systems Using Neural Networks", IEEE, 2015, Pp 497-501.

[3]. D. P. Gaikwad and Ravindra C. Thool "Intrusion Detection System Using Bagging Ensemble Method of Machine Learning", IEEE, 2015, Pp 291-295.

[4]. Shelly Xiaonan Wu and Wolfgang Banzhaf "The Use of Computational Intelligence in Intrusion Detection Systems: A Review", Applied Soft Computing, 2010, Pp 2-42.

[5]. Shi-Jinn Horng, Ming-Yang Su, Yuan-Hsin Chen, Tzong-Wann Kao, Rong-Jian Chen, Jui-Lin Lai and Citra Dwi Perkasa "A novel intrusion detection system based on hierarchical clustering and support vector machines", Elsevier, 2011, Pp 306-313.

[6]. Asaf Shabtai, Uri Kanonov and Yuval Elovici "Intrusion detection for mobile devices using the knowledge-based, temporal abstraction method", The Journal of Systems and Software, 2010, Pp 1524–1537.

[7]. Monowar H. Bhuyan, D. K. Bhattacharyya and J. K. Kalita "Network Anomaly Detection: Methods, Systems and Tools", IEEE, 2014, Pp 303-336.

[8]. S. Revathi and Dr A. Malathi "A Detailed Analysis on NSL-KDD Dataset Using Various Machine Learning Techniques for Intrusion Detection", IJERT., 2013, Pp 1848-1853.

[9]. Mahbod Tavallaee, Natalia Stakhanova and Ali A. Ghorbani "Towards Credible Evaluation of Anomaly-based Intrusion Detection Methods", IEEE, 2010, Pp 1-10.

[10]. Shaik Akbar, Dr.K.Nageswara Rao and Dr J.A.Chandulal "Intrusion Detection System Methodologies Based on Data Analysis", International Journal of Computer Applications, 2010, Pp 10-20.

[11]. Jayveer Singh and Manisha J. Nene "A Survey on Machine Learning Techniques for Intrusion Detection Systems", International Journal of Advanced Research in Computer and Communication Engineering, 2013, Pp 4349-4355.

[12]. A. M. Chandrashekhar and K. Raghuveer "Fortification of Hybrid Intrusion Detection System Using Variants of Neural Networks and Support Vector", IJNSA, 2013, Pp 71-90.

[13]. Álvaro Herrero, Martí Navarro, Emilio Corchado and Julián "RT-MOVICAB-IDS: Addressing Real-Time Intrusion Detection", Elsevier, 2013, Pp 1-24.

[14]. Bharanidharan Shanmugam and NorbikBashah Idris "Hybrid Intrusion Detection Systems (HIDS) using Fuzzy Logic", Intrusion Detection Systems, 2011, Pp 1-21.

[15]. Kumar, Vikash, Ditipriya Sinha, Ayan Kumar Das, Subhash Chandra Pandey, and Radha Tamal Goswami. "An integrated rule-based intrusion detection system: Analysis on UNSW-NB15 data set and the real-time online dataset." Cluster Computing (2019): 1-22.

[16]. Anshul Chaturvedi and Vineet Richharia, "A Novel Method for Intrusion Detection Based on SARSA and Radial Bias Feed Forward Network (RBFFN)", international journal of computers & technology, vol 7, no 3.

[17]. Jain, Upendra "An Efficient intrusion detection based on Decision Tree Classifier using feature Reduction", International Journal of Scientific and Research Publications, Vol. 2, Jan. 2012.

[18]. E. Blanzieri and A. Bryl "A survey of learning-based techniques of email spam filtering" Artif. Intell. Rev., vol. 29, no. 1, pp. 63–92, 2008.

[19]. D. Sculley and G. Cormack "Filtering email spam in the presence of noisy user feedback" in Proc. 5th Email Anti-Spam Conf., 2008, pp. 1– 10.

[20]. HengjieLi, Jiankun Wang "Intrusion Detection System by Integrating PCNN and Online Robust SVM" IFIP International Conference on Network and Parallel Computing, 2007.pp 250-255.

[21]. V. Engen, J. Vincent, and K. Phalp "Enhancing network-based intrusion detection for imbalanced data" Int. J. Knowl.-Based Intell. Eng. Syst., vol. 12, no. 5–6, pp. 357–367,2008.

[22]. S. T. Powers and J. He "A hybrid artificial immune system and self-organizing map for network intrusion detection" Inf. Sci., vol. 178, no. 15, pp. 3024–3042,2008.

[23]. K. Shafi, T. Kovacs, H. A. Abbass, and W. Zhu "Intrusion detection with evolutionary learning classifier systems" Nat. Comput., vol. 8, no. 1, pp. 3–27,2009.

[24]. Y. Yang and S. A. Elfayoumy "Anti-spam filtering using neural networks and Bayesian classifiers" in Proc. IEEE Int. Symp. Comput. Intell. Robot. Autom., 2007, pp.272–278.

[25]. Mohammad Behdad, Luigi Barone, Mohammed Bennamoun and Tim French "Nature-Inspired Techniques in the Context of Fraud Detection" in IEEE transactions on systems, man, and cybernetics part c: applications and reviews, vol. 42, no. 6, November 2012.

[26]. Mewada, Arvind, and Rupesh Kumar Dewang. "Deceptive reviewer detection by analyzing web data using HMM and similarity measures." Materials Today: Proceedings (2021).