Turkish Online Journal of Qualitative Inquiry (TOJQI)

Volume 12, Issue 3, June 2021: 2457-2467

Research Article

Constructing Security aware PSO based Approach for Preventing Cyber-attacks and Analysis using Healthcare Data

Karunanithi Senthamilselvi Vijayanand[1], D.Vinod[2]

Abstract

The cyber security key rules that the intelligence and techniques that has been taken cyber- deployment and massive data access all over the world. There are unpredicted issues and challenges that Artificial intelligence(AI) rapidly increases the threats also. Even though the security efficient provided that are incorporated along with machine learning and deep learning a major attacks are even eventually growing faster. Healthcare a foremost region that has to be available, integrity about patient's details and confidential security in data relies on cyber security. Whenever transaction and operation of healthcare data are manipulated and spread all over the research and distributed access the cyber-attacks involved too. The proposed work defines a context based cyber-security using Bio-stimulated cross approach artificial intelligence (CBSCA). The context combines the Deep learning approach method that is mostly applicable forcyber-attacks. Where the privacy and integrity of severe network applications that describes the health care analysis network.

Diversified approaches defined below

i) Diverse approach for anomaly detection model for protecting the data and security that can be shelter from cyber-attacks that cannot be evaded beyond security measures such as passive attacks. (EG: Antivirus)

ii)Learning ability monitoring for malware detection (LAMD) which fix and separate attacker's location in executable packets that cannot be undetectable by firewalls.

iii) Data manipulation on SQL preventing system attacks that has to be focused to avoid various SQL injection techniques.

The results show that the proposed system attains more accuracy than the existing works.

[1]Chief technology officer at Ja Secure Pte Ltd, Singapore
[2]Department of Computer Science and Engineering,
Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, India
[1]vj@ja-secure.com, [2]dvinopaul@gmail.com
Corresponding Author: dvinopaul@gmail.com

Keywords: AI, LAMD, CBSCA, SQL Injection, Bio-stimulated, Deep Learning .

1. Introduction

Protection of healthcare data is much necessary in the field of medicine so as to protect at the time of crisis. Maintaining the secrecy of the patient data along with confidentiality of these systems involved in management is a critical factor to be noted. In case of any loss of data, system should be designed in such a way to handle the critical situations. The techniques usually smart in nature are combined together in order to prevent attacks. The attacks methods that try to hack the sensitive data may be anyone as below,

1. Using attacks such as denial of service attacks that allow invasion of the system directly,

2. Installing software that are supposed to be malware.

3. Recognition of system weakness in case of security with the help of SQL injection attacks.

1.1 Using attacks such as denial of service attacks

If the attack happens directly, either firewall or IDS termed as Intrusion Detection Systems can be used to provide security. IDS are mainly associated with the monitoring of intrusion activities either by the users who may be authorized or even by other users. The events are analyzed and protection is given based on it. IDS is further divided into two types as the one based host as well the another one based on the network. The approaches employed for IDS building are as shown in Figure.1.

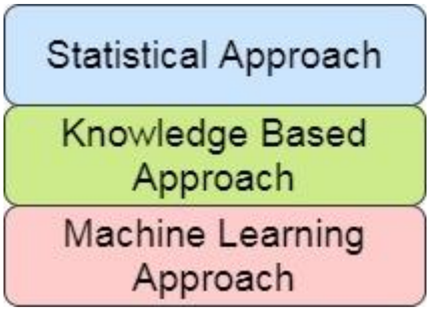

Figure.1. Approaches in IDS Building

The approaches that are based on statistics at initial stages are employed in detecting the activity of network that is considered to be normal. The other traffic that comes out of this normal scope are called to be anomalous in nature. The detection is accurate unless the system remains to be active in the network. The second type based on the knowledge work on classification of vectors by using the rule set or from the models which we have obtained based on the previous experience. Automation in data vectors analysis is carried out in the last approach called as machine learning. Also the systems which implement them also show a better performance.

1.2 Installing software that are supposed to be malware.

Antivirus software deployed systems will not be able to detect the malwares at all the times. A software that affects the working of the computer, collect sensitive data or attack the systems. Binary file in the system is searched for signatures that exist as predefined in the existing works. Although the antivirus scanner can be used for filtration of such files, programs that are supposed to be malicious bypass these. The system of the victim gets affected due to their execution. One of prominent methods to conceal the code that is supposed to malicious is called code packing. At the end of data processing, the code that is executable is converted as data by the developers of malware. Here analysis is done in a static way whereas during the runtime, data gets to its original form with the help of routine to restore. This stands as a point of entry for the actual malware after which the transformation takes place. The result is transparent execution and the control gets transferred to next layer.

Malicious contents need to be identified to find whether the unknown ones are available during packing. This is mainly used in the process of analysis of the original ones and this unpacking can also result in higher overhead while computation as well as requirement of resources. The time taken for processing varies which stands as a obstacle for detection of virus as large file scanning takes more time

1.3 Recognition of system weakness in case of security with the help of SQL injection attacks

The vulnerabilities related to the network applications can be exploited with this approach. One of the most unique code injection method where the query in SQL is inserted through the client input. This results in reading of data, modifying the data available in database and also execution of operations that are related to administration of database and so on. These attacks affect the normal execution of the SQL commands in the database and can be overcome by the proposed system

The paper is organized as follows, Section 2 deals with the existing systems available, Section 3 explains the proposed work, Section 4 presents the results obtained from the existing works and Section 5 concludes the paper.

2. Literature Review

Algorithms in domains such as artificial intelligence as well as data mining are used to identify the intrusion in systems by using methods such as clustering as well as classification. The former is a unsupervised learning method [1-3] while the latter is a supervised learning method [4-7]. Hybrid methods including neural network as well as neural networks are used along with genetic algorithms and many other systems and so on. In case of unpacking methods, dynamic approach was employed for the purpose of monitoring binary execution. This helps in extraction of the code which is original one. The above approach executes the code in environments that remain isolated such as a emulator as well as a virtual machine.

In case of any event happening, the execution gets stopped after tracing it. Heuristics can also be applied in order to identify the points where the jumps in the execution takes place. At this juncture, the content of memory becomes bulk. These approaches are found to be more difficult as well as consume more time and are also not able to counter the process of unpacking. Structural information can also be used in case of executable code for the purpose of training the classifiers which are supervised in nature. This helps in identifying whether the sample chosen is malicious in nature or not. They also utilized filtering method as a result of which the process consumes more time and also has more overhead in relation with the computation.

In order to detect the malicious activities both AI as well as data mining methods can be employed. They also help in discovery of the patterns of the malware. Compared to Naive Bayes method as well as SVM (Support Vector Machine), decision trees can be used to obtain effective results. Association rules can also be employed in the process of extraction in an automatic manner that differentiates the infected files from the ones that are not infected. Also Hidden Markov model was also made in use to identify whether the file is infected or not. These models have also been employed in field such as bioinformatics also. ANN termed as Artificial Neural Networks are also employed in detecting malware that are polymorphic in nature. Maps which are self-organizing in nature can also identify the malware patterns. But when it comes to accuracy, it seems to be low and also the malwares cannot be completely detected due to the false alarms also.

The attacks such as SQL injection can also be detected using analysis methods that are static in nature. But these approaches were not able to detect them completely prior to the beginning of the system. The proposed system also aimed at identifying the attacks before the time of execution by finding the queries that seemed to be illegal. Also filtering mechanism can be applied with the static methods of detection. Parse tree comparison with the SQL queries and its associated statements were also proposed. While using a web application, its internal state is being analyzed along with the relationships of the internal state of the application.

The SQL statements that are anomalous in nature have certain sections of code that is injected and are also found to be different from the other statements related with the applications. Both intrusion detection methods and machine learning models were employed to find the queries that missed matching during the comparison stages. Classification method in machine learning can be employed to identify the malicious code. In order to distinguish the malicious statements from the original ones, queries were represented in form of a tree. SVM can also be employed for classifying as well as predicting these attacks. Here again the comparison takes place to find the attacks. Bayesian classifier employed to detect the malicious code also made use of the HTTP parameters and mainly was dependent on the dataset quality employed for training. Pattern classifiers were employed for the purpose of identification of attacks to protect the applications. Conversion of these requests of HTTP into numeric attributes is performed in the proposed system. Parameter length along with the keyword count are the parameters taken into consideration. Based on these, classification is performed to find the patterns of injection in the parameters.

3. Proposed Work

The proposed work that has AI approach which has developing neural network using Radial basis function by fully connected ANN,FFNBP and PSO optimizing approach with genetic algorithm. In healthcare analysis security framework are major effective information’s that has to monitor without any attacks and how efficient by making less convention of possessions.

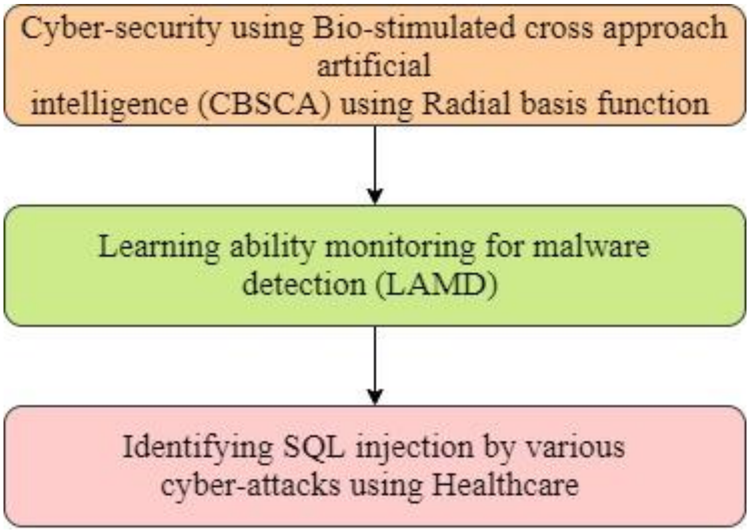
Cyber-security using Bio-stimulated cross approach artificial intelligence (CBSCA) using Radial basis function
Learning ability monitoring for malware detection (LAMD)
Identifying SQL injection by various cyber-attacks using Healthcare

Figure.2. Proposed System Functionality

3.1Cyber-security using Bio-stimulated cross approach artificial intelligence (CBSCA) using Radial basis function

Based on the three layer connected as fully connected layers such as neuron communication that has several steps to maintain the cyber-attacks removal. According to the ranking value for encoding the positive and negative Convolutional way for computation process neuron has weight and bias to modify the mapping communication. Based on the activity of learning pattern the multilayer perceptron the order ranking happened. Health care analysis data are impressively most necessary data for every part of data accessed. When the features that are required and classified as the sensible data variety as diseases rate, age, gender etc. In learning pass there is one entry learning where the neuron which has already trained that are in medical repository that has individual activation function named logistic sigmoid function. There are classes in trained model which features have propagated according to the classification processed from various feature analysis.

Classes can be defined in neuron as Cls= C1,C2,C3...Cn, weight based on neuron as We1 = W1,W2,W3..Wn and bias function bi= b1, b2..bn where the maximum order from the trained dataset from the distance that are calculated from the weight factors that are computed by using the rank order which have certain threshold value which have optimal solution according to the stored repository that are stored in neuron that can be called as output neuron.as mathematical modeling are basic part in ANN the activation function called the logistic stochastic function which have radial basis function.

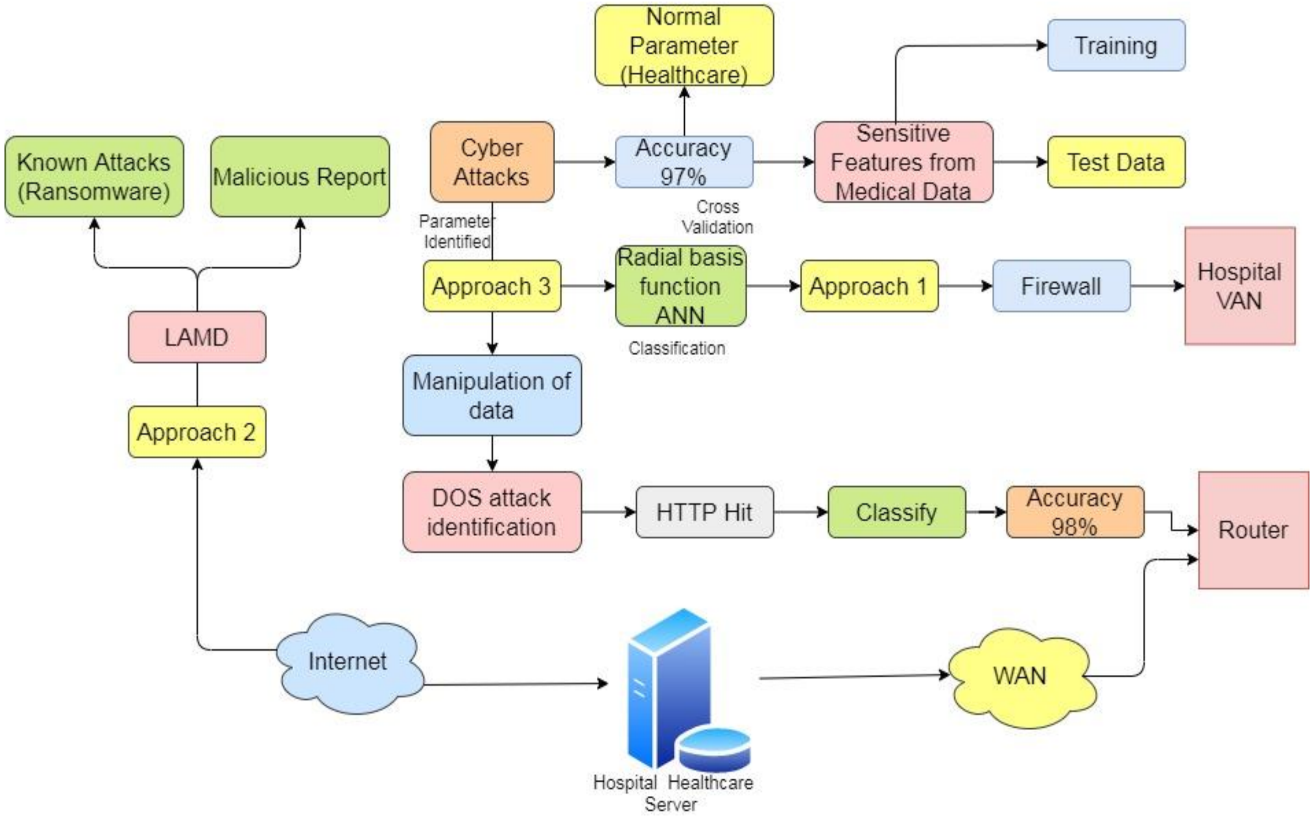

Figure.3.: Proposed work approach for Healthcare data classification and prediction analysis

The structure of parameters combination the approximation and accuracy can be derived by predicting the features.Corresponding the neuron insisted the hidden,input and output are being vector radically symmetric and communication based on the heart rate parameters of healthcare data the hidden layer connected to previous input are also recognized. These identical patterns required for maintain the distance reached according to the Gaussian value.

3.2 Learning ability monitoring for malware detection (LAMD)

A set of parameters from healthcare analysis based on the data operation differs in learning and computation methodologies.The flow of information based on the operation if database such as login credentials, level of dosage, medical expense from server side, time for classifying and the testing the subject that are modified according to the performance on data and time taken to access the data. As deep learning(DL) have data knowledge discovery depending on the membership relationship function which are depending on previous relation are generated. According to the rules that are associated as data are grouped based on learning as supervised, unsupervised. Thus the output features are Weight based coordination when the location of groups derived as cluster according to the equality of relations between each attributes.

Dataset curtsey from Canadian institute of cyber security (CIC) that has various anomaly detection and attacks categorized based on the amount of testing,classification and evaluated results are grouped. When the analogous behavior are identified from the shared information from the healthcare data and privacy characteristics are marked untraceable mapping according to the cluster that matched the pattern. LAMD an important feature mapping function that computes the connection based on the bias and model that are node applied to the mapped clusters. The prediction at the cluster data are applied with the example as below.

Gaussian function that have two analysis If,ELSE that maps the function whose pattern are maximum to the standard deviation value corresponding to the value expected from the value of the number of total clusters.

3.3 Identifying SQL injection by various cyber-attacks using Healthcare

MaliciousDoH dataset with 31 characterized that have traffic features, basedon time features, time-stamping, intrusion host based features suchas Denial of service, active attacks, passive attacks, malware abnormal datasets which have trained dataset which has 78% accuracy with 10,203 records and testing accuracy of 22% for testing results.

a) Process 1

When there are passive attacks oriented data even antivirus cannot restrict is being class labeled according to the multiclass membership variables. The min and max interval process and validated according to the manipulation. When the wireless LAN are connected to the healthcare analysis server which hits the http through the class attributes and resources are processed according to the radial basis function. The predictions according to the malicious repeated attacks are generated as report and then accuracy is also analyzed with its loss function.

b) Process 2

Based on the neural network computation process are trained and tested and applied according to the result analyzed and logistics sigmoid deviates the performance and submit the form from the server according to the access through the point on virtual private network processing both the admin side and user progress also.

c) Process 3

Class label as normal and abnormal which have features cluster as 5 features that are grouped based on the threshold value. When the vectors are classified and given to various hidden neural networks that can applied the parameter verification. The sensible data are given an alert signal when there is intrusion and cyber-attacks when the anonymous process are identified through the cross validation.

3.4 PSO - Particle Swarm Optimization

Particle Swam Optimization is one of the most important swam intelligent algorithms and performs better than the evolutionary algorithms. The former algorithms focus on betterment of the population at every iteration whereas the later one generates new population for each iteration.

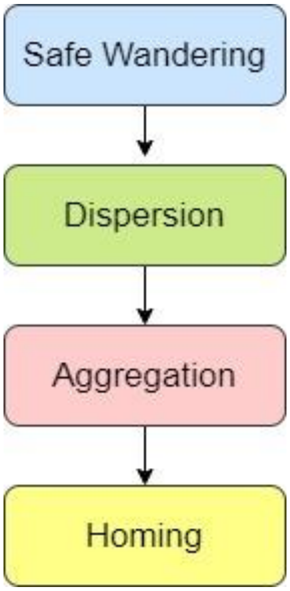

Figure.4. Bird Flocking rules

The basic PSO model does not follow the first two rules of the bird flocking rules. The last two rules are followed in the model. Swarm intelligence along with search space is used in PSO and random solutions that are potential are generated. These collection make up a swarm and every particle constitute the swarm. In PSO, the learning can either be social type where the others are involved for learning whereas the second one is said to be cognitive type. The best solution associated with the former type is gbest and the one associated with the later is pbest. Velocity of a particle decides the change regarding direction as well as magnitude. It gets updated based on the equation,

$$v_{id}^{t+1} = v_{id}^{t} + c_1 r_1 (p_{id}^{t} - x_{id}^{t}) + c_2 r_2 (p_{gd}^{t} - x_{id}^{t}) \quad \text{Eq.(1)}$$

When the velocity gets updated, position of the particle also gets updated.

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \quad \text{Eq.(2)}$$

3.4.1 Algorithm: PSO

```
$D_i$ - Dimensions, $S_i$ - Swarm Size, $p_i$ - particle index
Swarm S, Velocity vector initialized
for x = 1 to maximum iterations do
    for $p_i$ = 1 to Si
        for $D_i$ = 1 to d
            Perform update equation for velocity;
            Perform update equation for position;
        end
        Calculate updated position fitness value;
```

```
            Update gbest, pbest;
        end
        If gbest got best solution, terminate the problem;
End
```

4. Results and Discussion

In general the diverse approach for anomaly detection model using artificial neural networks where the healthcare data can be protected and safeguard from various type of attacks such as spoofing,data modification etc. The Feed forward neural network which has confidential health care data that may happened with malicious attackers who are actively available in server that are controlling the integral data. These issues from our LAMD dataset for intrusion detection systems that domain,IP address and patients control through port from admin and doctors who can anonymously enter and access the data. The healthcare data which has passive attacks and ransom aware attacks are actively monitored from various threats and monitoring are visualized based on the accuracy.

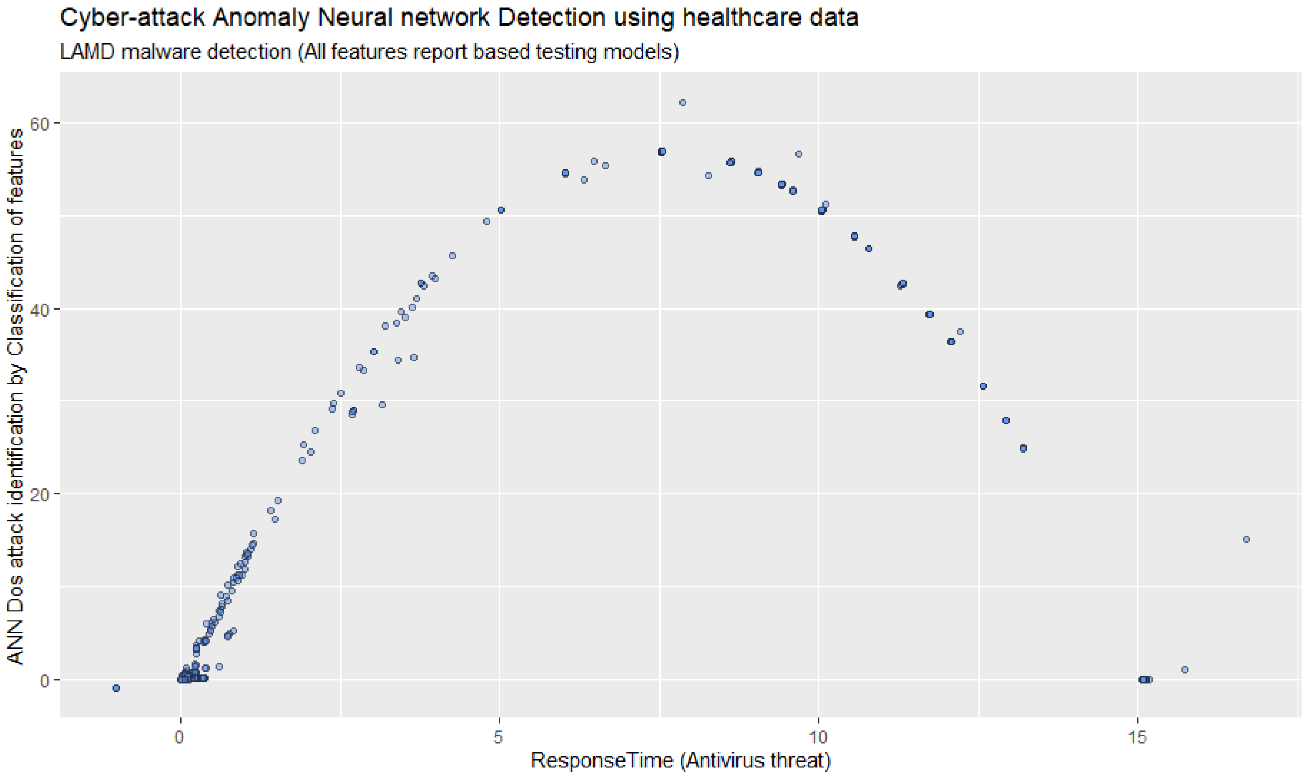

Figure 5. Cyber-attack Anomaly Neural Network Detection

The proposed method LAMD Learning ability monitoring for malware detection is reliable with approach 1 results which have 97% of accuracy that have visualized below.

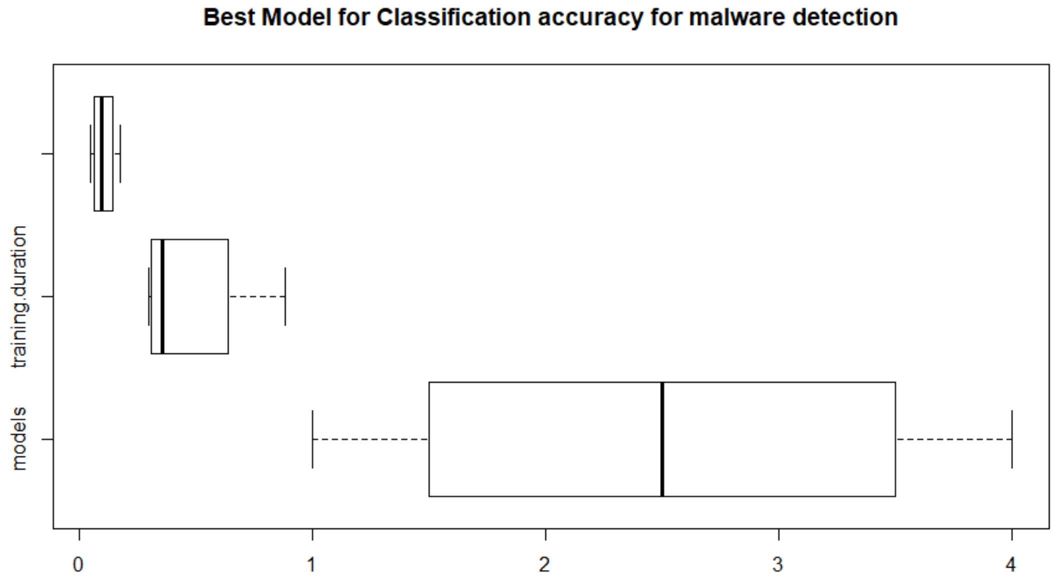

Figure 6. Performance Analysis of Proposed Work

Based on performance, the proposed system employing ANN shows more accurate results than the existing works such as SVM , KNN and so on. In case of training, the proposed system needs only less time for training. When it comes to testing, the systems using SVM and KNN require more time whereas the proposed work based on ANN needs only much less time.

5. Conclusion

Each system test based on artificial neural network that has multiple approach with various classification that shows the accuracy based on our dataset that has ransom ware attacks, Dos attacks etc. in our health care analysis for very efficient results. The more generalized LAMD approach have learning ability to detect the critical malware through the radial basis function to share the sensible data via a positive and negative points that has very good accurate according to the response time through the training and testing results. More over TP,FP via the precision F-measures are depending on the class features which has the selection for validation and optimized by choosing the prey using PSO optimization algorithm. By concluding the methods of AI better results are observed using the bio-stimulated cross approach.

References

1. Muna, M., Jawhar, T., Monica, M.: Design network intrusion system using hybrid fuzzy neuralnetwork. Int. J. Comput. Sci. Secur. 4(3), 285–294 (2009)
2. Jakir, H., Rahman, A., Sayeed, S., Samsuddin, K., Rokhani, F.: A modified hybrid fuzzyclustering algorithm for data partitions. Aust. J. Basic Appl. Sci. 5, 674–681 (2011)
3. Suguna, J., Selvi, A.M.: Ensemble fuzzy clustering for mixed numeric and categorical data.Int. J. Comput. Appl. 42, 19–23 (2012). doi:10.5120/5673-7705
4. Vladimir, V.: The Nature of Statistical Learning Theory, 2nd edn., p. 188. Springer, New York(1995). ISBN-10: 0387945598

5. John, G.H.: Estimating continuous distributions in bayesian classifiers. In: Proceedings of theEleventh Conference on Uncertainty in Artificial Intelligence, (UAI’ 95), pp. 338–345. MorganKaufmann Publishers Inc., San Francisco (1995)

6. Sang-Jun, H., Sung-Bae, C.: Evolutionary neural networks for anomaly detection basedon the behavior of a program. IEEE Trans. Syst. Man Cybern. 36, 559–570 (2005)doi:10.1109/TSMCB.2005.860136

7. Mehdi, M., Mohammad, Z.: A neural network based system for intrusion detection andclassification of attacks. In: IEEE International Conference on Advances in Intelligent Systems- Theory and Applications (2004)

8. Royal, P., Halpin, M., Dagon, D., Edmonds, R.: Polyunpack: automating the hidden-codeextraction of unpack-executing malware. In: ACSAC (2006)

9. Kang, M., Poosankam, P., Yin, H.: Renovo: a hidden code extractor for packed executables. In:2007 ACM Workshop on Recurring Malcode (2007)

10. Martignoni, L., Christodorescu, M., Jha, S.: Omniunpack: fast, generic, and safe unpacking ofmalware. In: Proceedings of the ACSAC, pp. 431/441 (2007)

11. Yegneswaran, V., Saidi, H., Porras, P., Sharif, M.: Eureka: a framework for enabling staticanalysis on malware. Technical Report SRI-CSL-08-01 (2008)

12. Danielescu, A.: Anti-debugging and anti-emulation techniques. Code-Breakers J. 5(1), 27–30(2008)

13. Farooq, M.: PE-Miner: mining structural information to detect malicious executables inrealtime. In: 12th Symposium on Recent Advances in ID, pp. 121–141. Springer, New York(2009)

14. Shaq, M., Tabish, S., Farooq, M.: PE-probe: leveraging packer detection and structuralinformation to detect malicious portable executables. In: Proceedings of the Virus BulletinConference (2009)

15. Perdisci, R., Lanzi, A., Lee, W.: McBoost: boosting scalability in malware collection andanalysis using statistical classification of executables. In: Proceedings of the 2008 AnnualComputer Security Applications Conference, pp. 301/310 (2008). ISSN: 1063–9527

16. Kolter, J.Z., Maloof, M.A.: Learning to detect and classify malicious executables in the wild.J. ML Res. 7, 2721–2744 (2006)

17. Ugarte-Pedrero, X., Santos, I., Bringas, P.G., Gastesi , M., Esparza, J.M.: Semi-supervisedLearning for Packed Executable Detection, Network and System Security (NSS), 5th InternationalConference on, (2011). DOI: 10.1109/ICNSS.2011.6060027

18. Ugarte-Pedrero, X., Santos, I., Laorden, C., Sanz, B., Bringas, G.P.: Collective classificationfor packed executable identification. In: ACM CEAS (2011)