Attack Detection within Network using Hybridization of Bisection TDE

Amit Dogra, Taqdir Abstract

Abstract: Today the leading cause of attack is traffic distribution that affects transmission as well as network performance. Traditionally the mechanism is utilized for detecting traffic distribution is not so accurate, user friendly and also they are time consuming. Therefore in this paper a hybrid approach is used that detect attacks by eliminating replicated features. It utilizes TDE with bisection mechanism that has great attack detection rate. The result of the proposed system as compared to existing system is better by the margin of 8 to 10%. The primary reason for the betterment is because of contrast enhancement and scaling factor at pre-processing and objective function overhead consideration at segmentation and classification phase. The results are shown in terms of accuracy, PSNR and MSE.

Keyword: TDE method, PSNR, traffic distribution.

I. INTRODUCTION

The main control centre of network is bandwidth that manage the overall activities of network like movement, memory, traffic etc. There are various disabilities that affect bandwidth activities. These are known as traffic or bulk attack. The attacks damage the bandwidth and also affect the network’s memory and other daily life’s activity. The various diagnostic methods are used for detecting traffic distribution and it utilizes mechanism to improve attack detection. But the current methods are not user friendly and invasive so there is need to develop convenient, non-invasive and user friendly detection and diagnostic system of traffic distribution. The Traditional differential technique is utilized for this purpose. Data segmentation is the process in which data is partitioned into multiple regions which results into number of fragments of objects. The main purpose of data segmentation is to make modifications and simplifications in data to make it more clear and meaningful. The main feature of segmentation is to find boundaries and objects in data. Resultant of data segmentation composed of multiple regions that entirely cover the whole data. Data in a common area are alike to each other on the basis of some computed property such as common difference, frequency. It is computer based process of bringing out useful information from the data. Input is the data and process ends with an output that is data which is numerical. Data processing is a technique by which we can change the input data by applying mathematical algorithms to produce an output data that is enhanced.
Attack Detection within Network using Hybridization of Bisection TDE

M.Mary Synthuja Jain Preetha et.al, presented[5], an algorithm for common difference data segmentation that is totally based on region growing. Basically it has phases, In first phase selection on initial seeds are done by having analysis of histogram than regions are allowed to grown from these selected seeds. In next phase threshold is optimized again but the use of cuckoo search algorithm. At last results of both has been compared against various evaluation parameters.

Xiaolin Zhang, Tao Yang et.al, 2015 presented[6], the limitations of older technique that is 2d entropy threshold segmentation. The researcher has proposed a flame data segmentation method that utilize artificial bee colony algorithm having the properties of levy flights. By utilizing strategy of levy flight id creates a recently created space of local search and increase the convergence rate. Computational results provide us with a methodology that is powerful in means of threshold selection correctness and reduced run time.

III. PROPOSED METHODOLOGY

In the proposed system we are using two classifiers namely bounding box and Haar like mechanism. The bounding box is used for detecting the region of interest and then Haar Like mechanism is used to detect feature from region of interest.

A. Bisection TDE Mechanism

In this mechanism the box surrounds the data in the training data with K dimensional and then compute the vector V using various frequency in that box.

\[ y = \beta + \sum_{i} \alpha_i (V, V_i) \]  

Here \( y \) is coordinate, \( \beta, \alpha \) are constant values, \( V \) is vector for various frequency

B. Haar Mechanism

In this the data is firstly decomposed using 2D Haar mechanism in multiple levels. The process produces four coefficients, i.e. approximation, horizontal, vertical and diagonal coefficients. Among these four coefficients, only approximation coefficient is chosen to generate feature vectors. The equations are as given below:

\[ L_k(M) = \iint K(M-x)I(x)dx \]  

Here \( K \) represents coefficient, \( M \) denotes the chosen feature, \( I(x) \) represents the number of feature vector.

Flowchart
A. Algorithm

Training
- Perform training using Haralick and Bisection mechanism from training data. (These features define region of interest (ROI) that are inputted by the user looking for an object)

Testing
- First of all, data are inputted for analysis must be converted into data frames. The standardized library of MATLAB for signal analysis can be used for this purpose. Mnread function is available in MATLAB and format of processing is .csv.
- After extracting the data, each data is labelled as Data1, Data2, Data3, ..., DataN
- Loop from first data to last data is applied for processing each data frame
  a) Extracting features using the Haralick feature extraction technique and store it in .mat file (Assuming frequency features of traffic data is different)
  b) Multithresholded bisection mechanism can be used to estimate the height of the object from the data
  c) The extracted feature corresponding to height will be maintained within .mat file.
  d) If (Trainingg ≠ Testingg) then Object detected Return Else Frame=Frame+1 End of if

End of Loop
- Stop

This methodology correctly classifies leaf into distinct categories. The region of interest is increased through the H-Bisection TDE classifier and hence more region is classifier and non-segmented region is reduced. Moreover, proposed approach is dynamic since any dataset can be analyzed rather than fixed set of data hence applications of both supervised and unsupervised learning.

IV. RESULTS

The result obtained after applying bisective TDE mechanism is in terms of PSNR and MSE. In addition classification accuracy is also considered for proving validity of study. This metric must be enhanced and indicates correctly identified segments from the data within the dataset. Classification accuracy is expressed in the form of true positive, true negative, false positive and false negative parameters as in the equation 3

\[ \text{Classification Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \]

Equation 8

TP is the true positive and TN is the true negative values obtained during simulation.

<table>
<thead>
<tr>
<th>Table 1: Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Used Data</strong></td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>1000 tuples</td>
</tr>
<tr>
<td>2000 tuples</td>
</tr>
<tr>
<td>3000 tuples</td>
</tr>
</tbody>
</table>

Classification accuracy is identified in terms of correctly classified segments from the data dataset. Result is expressed in the form of true positive rate. True negative rate is not observed while considering this parameter and hence classes identified may vary with time. In other words, classification accuracy for all the categories may not be symmetrical.

The peak signal to noise ratio obtained from the proposed system is better by 2 to 3%. The existing system PSNR is in the range of 12-13 where as proposed system peak signal to noise ratio is between 16-21. The peak signal to noise ratio must be higher for higher convergence rate. The evaluation equation for peak signal to noise ratio is given as

\[ \text{PSNR} = 10 \log_{10} \left( \frac{MSE}{255} \right) \]

Equation 9: Peak signal to noise ratio

Peak signal to noise ratio is expressed as mean square error and 255 is the maximum intensity of pixel. Table 2 gives the peak signal to noise ratio of existing and proposed system.

<table>
<thead>
<tr>
<th>Table 2: PSNR of existing and proposed system</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Used Data</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>1000 tuples</td>
</tr>
<tr>
<td>2000 tuples</td>
</tr>
<tr>
<td>3000 tuples</td>
</tr>
</tbody>
</table>

The peak signal to noise ratio expressing the desired levels of accuracy from the proposed system is better as compared to existing system. The plot elaborating table 6.3.2 is given as The mean square error is another parameter on which classification accuracy is greatly dependent.
Peak signal to noise ratio goes side by side along with mean square error. Higher the PSNR least will be MSE. In other words, mean square error can be obtained from PSNR itself.

\[ MSE = \sum_{i=1}^{n} \frac{\log(X_i)}{255}/n \]

Equation 10: Mean square error evaluation

‘X’ is the intensity of the pixel present within the data. ‘n’ is the total number of data within the data. As the mean square error decreases in the proposed system since population decreases, hence execution time also decreases subsequently.

<table>
<thead>
<tr>
<th>Used Data</th>
<th>Classified as</th>
<th>TDE time</th>
<th>CPU time</th>
<th>TDE-B Classifier time</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000 tuples</td>
<td>Attack detected</td>
<td>1.3098</td>
<td>1.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000 tuples</td>
<td>Not Detected</td>
<td>1.4562</td>
<td>1.095</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3000 tuples</td>
<td>Attack Detected</td>
<td>1.8786</td>
<td>1.0083</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Execution time reduction is due to feature vector reduction and replicated features are eliminated from the simulation. The population size decreases and hence overall execution time is decreases significantly.

V. CONCLUSION

The hybrid approach of TDE with bisection mechanism used in the proposed system eliminate the replicated features and thus malicious data is detected at great rate. The result of the proposed system as compared to existing system is better by the margin of 8 to 10%. The primary reason for the betterment is because of contrast enhancement and scaling factor at pre-processing and objective function overhead consideration at segmentation and classification phase. Since we are working upon the networking field, so quick result generation is the need of the hour. This is achieved using the proposed system. Individual distinct features obtained can be used for real time data also for attack detection. In future this mechanism can be tested against the real time data and performance can be tested on GPU with deep learning.

REFERENCES


AUTHORS PROFILE

Dr. Taqdir, Designation : Assistant Professor Qualification : PhD(CSE,NIT Jalandhar),M.Tech(CSE.),B.Tech(CSE) Department : Computer Science And Engineering Email : taqdir_8@rediffmail.com Dr Taqdir is currently working as Assistant Professor in the department of Computer Science and Engineering at Guru Nanak Dev University, R/C Gurdaspur. She hold the experience of 17 years in teaching at University level and her area of interest are Digital Image Processing, Machine Learning.

Mr. Amit Dogra, Designation : Assistant Professor Qualification : M.Tech, B.E Department : Computer Science And Engineering Email : aamitdoga004@gmail.com Amit Dogra is currently working as Assistant Professor in the department of computer science and Engineering in SoET,BGSB University Rajouri (J&K). He hold B.E degree in Information Technology and M.Tech degree in Networking and Internet Engineering.