RESEARCHERID
THOMSON REUTERS

[Bist * *et al.*, 6(10): October, 2017]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A SOLUTION TOWARDS OVERLAPPING CLUSTERS OF MALWARES USING MKM-PSO ALGORITHM

Ankur Singh Bist*
*CSE Department, KIET, Ghaziabad

DOI: 10.5281/zenodo.1002634

ABSTRACT

Complexity of network applications has enhanced the problem of computer worms. At the same time the advancement in technology has given rise to evolution of portable mobile devices. It has become the danger point of day today activities like mails, mobile banking etc. Our purpose is to design an algorithm that can answer the problem of overlapping malware clusters.

KEYWORDS: Classifiers, Classification Methods, Computer Virus

I. INTRODUCTION

Internet has become target of malicious codes due to its increasing use. Malicious codes are executable code and have the capability to replicate. It makes their survival strong. Viruses design and evolution attached with the area of programming. Similar to other computer programs viruses carry functions that are intelligent for providing protection in such a manner that detection remains not easy for virus scanner [1].

Viruses have to take various procedures of intellect for continued existence. That is why they may have complex encrypting and decrypting engines. These are the most frequent methods used by computer viruses in current scenario. They make use of these techniques to mask the antivirus and to adopt the certain environment for their expansion [2].

Figure 1: Assembly code of Virus File

Polymorphic viruses try to hide the decrypting module. More complex methods were developed enabling the virus designers to change the code of one virus file and make multiple morphed copies while maintaining its functionalities. These are the type of viruses which have the ability to mutate itself with the code changed but without changing its functionalities. Metamorphic virus can become a serious threat considering the fact that there can be thousands of variants of one virus file with their signature being totally different.

Metamorphic viruses transform its code in a specific manner very frequently and require to be prohibited. Their analysis will lead to evolve a framework where the overall process of detection will be bounded in specific outcomes of continuing evolving results. It is essential to make a distinction between replicating programs and its similar forms. Reproducing programs will not necessarily damage your system [3] [4] [5]. There is big fight between designers of virus and antivirus. The enhanced knowledge about the certain patterns, specifications can be designed. Various malicious codes can be evolved and incremented in well precise and efficient manner. For

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

perfect identification of a metamorphic virus, identification routines must be written that can generate the essential instruction set of the virus code from the actual occurrence of the infection [6-12].

Code obfuscation is one of the important properties adopted by metamorphic viruses. The mutating behavior of metamorphic viruses is due to code obfuscation techniques .There are various code obfuscation techniques.

a. Dead code insertion
b. Variable renaming
c. Break and join transformation
d. Expressing reshaping
e. Statement reordering

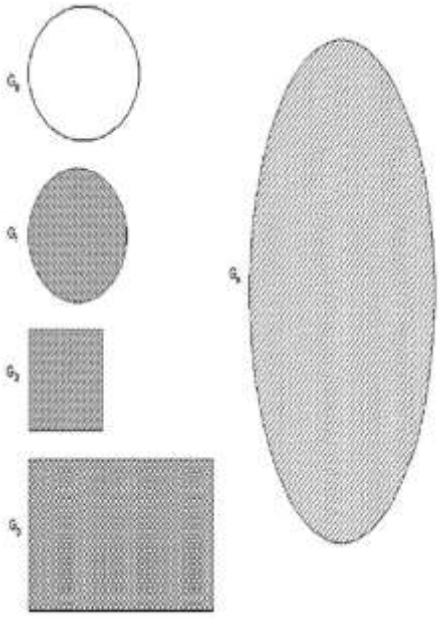

Figure 2: Analogy of Metamorphic Viruses

II. PROPOSED ALGORITHM (MODIFIED KMEANS-PSO) (MKM-PSO)

Malware samples are generated from NGVCK, MWOR kit and half of the samples are taken from different web links. Normal files are taken from windows. The score is generated with the help of pair-wise alignment algorithm.

III. K-MEANS ALGORITHM

Step 1: Input data set, clustering variable and maximum number of clusters
Step 2: Initialize cluster centroid
Step 3: Calculate Euclidean distance
Step 4: Move on to next observation and calculate Euclidean Distance
Step 5: Calculate Euclidean Distance for the next observation, assign next observation based on minimum euclidean distance and update the cluster centroids.

IV. PARTICLE SWARM OPTIMIZATION ALGORITHM

Generate random population of N solutions (particles);
For each individual i∈N: Calculate fitness(i);
Initialize the value of weight factor, w;
For each particle;

Set pBest as the best position of particle i;
If fitness(i) is better than pBest;
pBest(i)=ftness(i);
End;

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Set gBest as the best fitness of all particles;
For each particle;
Calculate particle velocity;
Update particle position;
End;

Update the value of weight factor,w;
Check if termination=true;
End;

[In this scenario PSO is used to optimize cluster center]

Code 1

Step1: Z_i=Objective data points. Ci, Cj are two clusters.
After clustering find(Zi)
Finds whether Zi belongs to Ci or Cj for all i
And calculate freq(Zi)
Finds frequency of data points getting overlapped.
Severe Function= freq(Zi)/(point(Ci)+point(Cj))

Step 2: Move cluster Ci and Cj in such a manner.
Move center of Ci to (-£) distance and relative move all data points strictly belongs to Ci and Cj to (£) until find(Zi)==Null

Step 3: while (Zi==0)
If(Hamm_distance(Zi)toCenter(Ci)>Hamm_distance(Zi)toCenter(Cj))
Ci←Zi
Zi - -
else
Cj←Zi
Zi - -
End if
End while

Following parameters are calculated:-

1. Accuracy=$\frac{TP+TN}{TP+TN+FP+FN}$
2. Precision=$\frac{TP}{TP+FP}$
3. Recall=$\frac{TP}{TP+FN}$
4. F-measure=$\frac{2*precision*recall}{precision+recall}$
5. ROC- Receiver Operating Characteristic (ROC) curve defines how a identification rate change as the internal threshold varies to produce more or fewer false alarm.

V. CONCLUSIONS

Large number of research papers has been written in the field of overlapping clusters. Our purpose in this paper is to address the problem of overlapping malware clusters. After experimentation it is observed that this process improved the classification accuracy. In future we will test the same algorithm for different data sets.

VI. ACKNOWLEDGEMENTS

This work was supported by Dept. of Computer Science and Engineering, Krishna Institute of Engineering and Technology (K.I.E.T) Ghaziabad, India.

VII. REFERENCES

1. Bist, Ankur Singh. "Classification and identification of Malicious codes." (2012).
2. Bist, Ankur Singh. "Hybrid model for Computer Viruses: an Approach towards Ideal Behavior." (2012).
3. Bist, Ankur Singh. "Detection of metamorphic viruses: A survey." *Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on*. IEEE, 2014.
4. Sharma, Rudranshu, and Ankur Singh Bist. "Genetic algorithm based weighted extreme learning machine for binary imbalance learning." *Cognitive Computing and Information Processing (CCIP), 2015 International Conference on*. IEEE, 2015.
5. Hartigan, John A., and Manchek A. Wong. "Algorithm AS 136: A k-means clustering algorithm." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28.1 (1979): 100-108.
6. Sharma, Rudranshu, and Ankur Singh Bist. "INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY MACHINE LEARNING: A SURVEY."
7. Das, Purushottam, and Ankur Singh Bist. "INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY FEATURE SELECTION."
8. Kumar, Babeesh, Sushila Vikas Maheshkar, and Ankur Singh Bist. "INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY Image Segmentation using Enhanced K-means clustering with divide and Conquer Approach."
9. Kumar, Vikas, and Ankur Singh Bist. "INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY CLUSTER ANALYSIS: A SURVEY."
10. Pandey, Neha, B. K. Singh, and Ankur Singh Bist. "A novel feature learning for image classification using wrapper approach in GA." *Signal Processing and Integrated Networks (SPIN), 2015 2nd International Conference on*. IEEE, 2015.
11. Sharma, Rudranshu, Ankur Singh Bist, and Vikas Kumar. "INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY EXTREME LEARNING MACHINE."
12. Shi, Y., & Eberhart, R. C. (1999). Empirical study of particle swarm optimization. In *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on* (Vol. 3, pp. 1945-1950). IEEE.

CITE AN ARTICLE

Bist, A. S. (2017). A SOLUTION TOWARDS OVERLAPPING CLUSTERS OF MALWARES USING MKM-PSO ALGORITHM .*INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, 6*(10), 62-65.