Abstract

A virus is defined as a program that spreads or replicates by copying itself, and generally has malicious effects. The antivirus systems used today mainly detect malware on the basis of known virus patterns, making detection of a new virus very difficult. This deficiency can be overcome by training an artificial neural network with the inputs from Portable Executable (PE) Structure of executable files, as they learn from the training data and will be able to identify unknown virus patterns. PE Structure contains various fields by which one can identify virus infected executable files from the legitimate ones without executing them, and Fisher Score can be used to select the most relevant features (fields) to speed up the analysis. A new technique of identifying virus infected files by using Fisher Score and applying them as input to the neural network is proposed.
Virus Detection using Artificial Neural Networks

References

- Dima Stopel, Zvi Boger, Robert Moskovitch, Yuval Shahar and Yuval Elovici, "Improving Worm Detection with Artificial Neural Networks through Feature Selection and Temporal Analysis Techniques.
- Quanquan Gu, Zhenhui Li, Jiawei Han, "Generalized Fisher Score for Feature Selection".
- Xiaofei He, Deng Cai, Partha Niyogi, "Laplacian Score for Feature Selection".
- Zheng Zhao, Lei Wang, Huan Liu, and Jieping Ye, "On Similarity Preserving Feature Selection".
- Daoqiang Zhang, Songcan Chen and Zhi-Hua Zhou, "Constraint Score: A New Filter Method for Feature Selection with Pairwise Constraints".
- Koji Tsuda, Motoaki Kawanabe and Klaus-Robert Muller, "Clustering with the Fisher Score".
- B. Azhagusundari, Antony Selvadoss Thanamani, "Feature Selection based on Information Gain".
- Rahman Mukras, Nirmalie Wiratunga, Robert Lothian, Sutanu Chakraborti, and David Harper, "Information Gain Feature Selection for Ordinal Text Classification using Probability Re-distribution".
- Zhaohui Zheng, Xiaoyun Wu, Rohini Srihari, "Feature Selection for Text Categorization on Imbalanced Data".
- Phayung Meesad, Pudsadee Boonrawd and Vatinee Nuipian, "A Chi-Square-Test for Word Importance Differentiation in Text Classification".
- George Forman, "An Extensive Empirical Study of Feature Selection Metrics for Text Classification".
- Anuj Sharma, Shubhamoy Dey, "Performance Investigation of Feature Selection Methods and Sentiment Lexicons for Sentiment Analysis".
- Todor Ganchev, Panagiotis Zervas, Nikos Fakotakis, George Kokkinakis, "Benchmarking Feature Selection Techniques on the Speaker Verification Task".
- Huan Liu and Lei Yu, "Toward Integrating Feature Selection Algorithms for"
Classification and Clustering”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 17, NO. 4, APRIL 2005.
- Eric Filiol, &quot;Malware pattern scanning schemes secure against black-box analysis”, Springer-Verlag France 2006.
- Yanfang Ye, Dingding Wang, Tao Li, Dongyi Ye, Qingshan Jiang, &quot;An intelligent PE-malware detection system based on association mining”, Springer-Verlag France 2008.
- Zheng Zhao, Fred Morstatter, Shashvata Sharma, Salem Aleyani, AneethAnand, Huan Liu, &quot;Advancing Feature Selection Research- ASU Feature Selection Repository&quot;.
- Yibin Liao, &quot;PE-Header-Based Malware Study and Detection&quot;.
- Joel Yonts, &quot;Attributes of Malicious Files&quot;, SANS Institute InfoSec Reading Room.

Index Terms

Computer Science  Security

Keywords

PE Structure  Feature  Fisher Score  Artificial Neural Network