Abstract

Intrusion detection systems intend to recognize attacks with a low false positive rate and high detection rate. Many feature selection methods introduced to eliminate redundant and irrelevant
features, because raw features may abbreviate accuracy or robustness of classification. In this paper we are proposing the information gain technique for the selection of the features. A feature with the highest information gain is the criteria for the selection of the features. We reduced the features of the data set than run the algorithm. Result show that drastically decreased in learning time of the algorithm without compromising the accuracy which is desirable for good IDS. We analyse two learning algorithms (NB and BayesNet) for the task of detecting intrusions and compare their relative performances. We comment on the suitability of the BayesNet algorithm for the intrusion detection task based on its high accuracy and high true positive rate. We finally state the usefulness of machine learning to the field of intrusion detection.

Reference

- Knowledge Discovery in Databases DARPA archive. Task Description. KDDCUP 1999 Data Set http://www.kdd.ics.uci.edu/databases/kddcup99/task.htm
- Pingchuan Ma,” Log Analysis-Based Intrusion Detection via Unsupervised Learning” Master of Science, School of Informatics, University of Edinburgh, 2003.

Intrusion Detection using Supervised Learning with Feature Set Reduction


- Shaohua Teng, Hongle Du Wei Zhang, Xiufen Fu “A Cooperative Network Intrusion Detection Based on Heterogeneous Distance Function Clustering” 2010, 14th IEEE International Conference on Computer Supported Cooperative Work in Design.

Index Terms

Computer Science

Security

Key words
<table>
<thead>
<tr>
<th>Intrusion Detection</th>
<th>Machine Learning</th>
<th>BayesNet</th>
</tr>
</thead>
<tbody>
<tr>
<td>KDD 99</td>
<td></td>
<td>NB</td>
</tr>
</tbody>
</table>