International Journal of Information Processing, 7(1), 51-56, 2013
ISSN : 0973-8215
IK International Publishing House Pvt. Ltd., New Delhi, India

Comparison of Supervised Learning and Reinforcement Learning in Intrusion Domain

Nandita Sengupta[a] and Jaya Sil[b]

[a]Information Technology Program, University College of Bahrain,
P.O. Box 55040, Manama, Kingdom of Bahrain, Contact: ngupta@ucb.edu.bh

[b]Department of Computer Science and Technology, Bengal Engineering and Science University
Shibpur, P.O. Botanic Garden, Howrah, Pin 711103, West Bengal, India, Contact: js@cs.becs.ac.in

In modern world use of network is increasing exponentially. Network security needs attention of computer science researchers. Intrusion Detection System is software / hardware which detects intruder in the network or host system. Classification plays an important role in Intrusion Detection System. Detection of anomaly or normal traffic is main working philosophy for such type of system. For detection of online traffic, learning of the system is required. In our paper, performance of Supervised Learning and Reinforcement Learning is compared in Intrusion Domain. NSL-KDD data is considered for our work. In that dataset for each object 41 conditional attributes and one decision class attribute are mentioned. Out of 41 attributes, 7 attributes are discrete and 34 attributes are continuous. Using feature ranking method, number of discrete attributes are reduced and these reduced number of attributes are used for classification in Supervised Learning. Some Supervised Learning like CS-MC4, Decision List, ID3, Naive Bayes, C4.5, Rnd Tree are applied on this data set and compared this classification result with classification accuracy derived from Reinforcement Learning combined with Rough Set Theory classifier.

Keywords: Classification, Intrusion Detection System, Reinforcement Learning, Supervised Learning.

1. INTRODUCTION

Online classification of network traffic data is very important to develop Intrusion Detection System (IDS) that automatically monitors the flow of network packets. Existing works on intrusion detection have been carried out to classify the network traffic as anomaly or normal. A majority of current IDS follow signature based approach [1] in which, similar to virus scanners, events are detected that match specific predefined patterns known as "signatures". The limitation of these signature-based IDS is their failure to identify novel attacks and even minor variation of patterns are not detected accurately. In addition, sometimes IDS generate false alarm for alerting network administrator due to failure of handling imprecise data which has high possibility to appear in network traffic data. Therefore, accuracy, computation time and system learning are the key issues to be addressed properly for classifying such data.

Classification is an important task in data mining research that facilitates analysis of huge amount of data. Learning plays an important role in classification for any dataset. A lot of research work on supervised learning, unsupervised learning has been carried out for intrusion detection system. Reinforcement Learning for intrusion detection system is an active research area where researchers can contribute to improve the performance of IDS. Feature reduction helps in optimization of classification with respect to time and efficiency. In the paper, network traffic data [2] of NSL-KDD has been considered for generating training and testing patterns. Feature ranking technology is used for feature reduction. In our work, 7 discrete attributes are considered and applying feature ranking technology, number

Sngapore in 2002-03 and visited Heidelburg University, Germany in 2007. Dr. Sil contributed a Book Chapter - Adaptive Agent Integration in Designing Object- Based Multiagent System. LNCS, Volume 3215/2004. Dr. Sil acts as Guest Editor In International Journal On Artificial Intelligence And Soft Computing And Editor of GA Issue Of Materials and Manufacturing Processes. Delivered Tutorial lectures in two Intl. Conferences NGMS 2006, 2008 and INDO US workshop in Kolkata. Her areas of research include Image Processing, Soft Computing Techniques, Multiagent Systems and Bio-Informatics.