Conceptual Design

Submitted by:

Group Number 26 [Manikanta Desu(203174002), Mayank Seth(203190001), Vivek Rathore(190100138), Ashara Keyur Ashwin(19D100003), Akash Rajendra Gayakwad(190260004)]

Detection Method:

There are two methods for detecting intrusion, Anomaly based and Signature based.

Anomaly Based	Signature Based
The model is trained using a training dataset and it also learns the behaviour of the system to detect intrusions.	Here the signatures of every connection are compared with signatures in a database to detect intrusions.
This method is less data intensive and hence faster.	This method is data intensive and hence slower.
It can detect unknown intrusions for which it was not trained.	It cannot detect intrusions which are not in the database.

We can infer that Anomaly based detection is more robust, reliable and faster than Signature based detection and hence we have used Anomaly based detection.

Dataset:

NSL-KDD dataset is used for training and testing the model for anomaly detection.

Classification Method:

In order to increase Detection Rate(DR), without increasing False Positive Rate(FPR), we have used multiple classifiers. We have used KNN and Random Forest classifiers to detect intrusions. We then combined the results of both of these classifiers using hard voting and a voting classifier. The result for the voting classifier was better than that for hard voting and hence we have used a voting classifier. Based on a threshold probability, a connection is either flagged as normal or threat. We have observed the variation of DR and FPR with threshold value and we have tried to achieve a higher DR which has led to a slight increase in FPR, but since we are now able to detect intrusions efficiently, slight increase in FPR can be neglected.