Introduction:  
Anomaly detection is the process of finding instances in a data set which are dierent from the majority of the data. It is used in a variety of application domains. In the network security domain it is referred to as intrusion detection, the process of finding outlying instances in network tra‑c or in system calls of computers indicating compromised systems. In the forensics domain, anomaly detection is also heavily used and known as outlier detection, fraud detection, misuse detection or behavioral analysis. Applications include the detection of payment fraud analyzing credit card transactions, the detection of business crime analyzing financial transactional data or the detection of data leaks from company servers in data leakage prevention (DLP) systems. Furthermore, anomaly detection has been applied in the medical domain as well by monitoring vital functions of patients and it is used for detecting failures in complex systems, for example during space shuttle launches.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sr No. | Research Papers and Publication | Algorithms | Procedure | Advantage | Limitations |
| **01** | | **CLASSIFIER BASED ANOMALY DETECTION METHODS:** | | | |
| 1 | Publisher : Liron Bergman Yedid Hoshen School of Computer Science and Engineering  Link :  https://openreview.net/pdf?id=H1lK\_lBtvS | GOAD  Training and Evaluation Algorithm | | The ability to generate any number of tasks.The anomaly detection performance on the KDD-Rev dataset with different numbers of tasks. From 16 tasks, the accuracy remains stable. | A small number of tasks (less than 16) leads to poor results. |
| 2 | Fuzzy Network Profiling for Intrusion Detection  Publisher: IEEE  Link:  https://ieeexplore.ieee.org/abstract/document/877441 | Fuzzy Intrusion Recognition Engine (FIRE) using fuzzy sets and fuzzy rules. | | When combined with data mining of input data to reduce the size of the input data sets and to select features that highlight anomalies, fuzzy logic can be an effective means of defining network attacks. | The datamining process reduces the amount of data that must be retained for historical comparisons of network activity, meaningful to anomaly detection than the raw input data. |
| 3 | Using an Ensemble of One-Class SVM Classifiers to Harden Payload-based Anomaly Detection Systems  Publisher: IEEE  Link : https://ieeexplore.ieee.org/abstract/document/4053075 | One-Class SVM Classifiers | | Particularly useful in case of two-class learning problems whereby one of the classes, referred to as target class, is well-sampled. | Particularly useful in case of two-class learning |
| 4 | MD Classifiers | | The MD classifier performs extremely well for ν = 0 and k = 10. In this case the MD classifier is able to detect all of the 18 attacks for an RFP around 0.1% and reaches 100% of detection rate for an RFP around 1% | Use of one classifier does not improve the hardness of evasion |
| **02** | | **Nearest K neighbours Anomaly Detection** | | | |
| 5 | Statistical Analysis of Nearest Neighbor and Cluster Analysis Methods for Anomaly Detection  By: Xiaoyi Gu et- Department of Statistics and Data Science, CMU | Nearest k neighbours in the surrounding for global anomalies | For every record find k nearest neighbors and find the average distance from the dataset.Finally normalize the data. | Main Advantages:  1) A straightforward way for detecting anomalies.  2) Easy to perform and maintain over a range of data  3)NND descriptor provides a good discrimination between normal and anomalous events. | When the number of the datasets increases then the offered efficiency decreases significantly.  The user has to specify k (the number of clusters) in the beginning  k-means can only handle numerical data |
| **03** | | **Cluster Analysis** | | | |
| 6 | Unsupervised Anomaly Detection Using an  Optimized K-Nearest Neighbours Algorithm  By Michael J. Prerau, Eleazar Eskin | Cluster Analysis optimization for Nearest k neighbours | Cluster analysis is applied before applying the nearest k neighbour algorithm and finally the data is normalized and optimized. | The anomaly detection rate increased from 84%(without cluster analysis ) to 91%  The mix algo works better because it forms a two layer mesh to detect an anomaly. | k-means assumes that we deal with spherical clusters and that each cluster has roughly equal numbers of observations |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **04** | **Statistical methods for Anomaly detection** | | | |
| 7) HBOS | Histogram-based Outlier Score (HBOS): A fast Unsupervised Anomaly Detection Algorithm Markus Goldstein and Andreas Dengel German Research Center for Artificial Intelligence (DFKI), Trippstadter Str. 122, 67663 Kaiserslautern, Germany  Histogram-Based Outlier Scores to Detect Computer Network Anomalies by Nerijus Paulauskas , et alDepartment of Computer Science and Communications Technologies, Vilnius Gediminas Technical | 1)univariate histogram is constructed first for every feature.  2)for categorical data, simple counting of the values of each category is performed and the relative frequency is found.  3) For numerical features, two methods can be used: (1) Static bin-width histograms (2) dynamic bin-width histograms.  4) The frequency of samples in each bin is used as an estimate of the density For dynamic bin-width: values are sorted first and then a mixed amount of N k successive values is grouped into a single bin where N is the number of total instances and k the number of bins. | HBOS was significantly faster than both. On larger data sets the speed-up can be much higher.  The higher number of bins reduces the influence of data grouping.  Dynamic bins shows better detection of dynamic events | Precision and Accuracy of the technique is less than other unsupervised methods |
| **04** | **Subspace based** | | | |
| 1. RPCA | Roland Kwitt and Ulrich Hofmann. 2007. Unsupervised anomaly detection in  network trac by means of robust PCA. In Computing in the Global Information  Technology, 2007. ICCGI 2007. International Multi-Conference on. IEEE, 37–37.  Roland Kwitt and Ulrich Hofmann. 2007. Unsupervised anomaly detection in  network trac by means of robust PCA. In Computing in the Global Information  Technology, 2007. ICCGI 2007. International Multi-Conference on. IEEE, 37–37.  Roland Kwitt and Ulrich Hofmann. 2007. Unsupervised anomaly detection in  network traffic by means of robust PCA. In Computing in the Global Information  Technology, 2007. ICCGI 2007. International Multi-Conference on. IEEE, 37–37. | Robust Principal Component Analysis (rPCA)  [  24  ] is based  on the Principal Component Analysis (PCA), that is used for dimen-  sionality reduction. PCA is used to detect subspaces in a dataset  and has been applied to anomaly detection to identify deviations  from the ’expected’ subspaces, which may indicate anomalous data  points. The principal components of PCA are the eigenvectors of the  covariance matrix, which is computed twice to improve robustness.  Robust Principal Component Analysis (rPCA)  [  24  ] is based  on the Principal Component Analysis (PCA), that is used for dimen-  sionality reduction. PCA is used to detect subspaces in a dataset  and has been applied to anomaly detection to identify deviations  from the ’expected’ subspaces, which may indicate anomalous data  points. The principal components of PCA are the eigenvectors of the  covariance matrix, which is computed twice to improve robustness.  Robust Principal Component Analysis (rPCA)  [  24  ] is based  on the Principal Component Analysis (PCA), that is used for dimen-  sionality reduction. PCA is used to detect subspaces in a dataset  and has been applied to anomaly detection to identify deviations  from the ’expected’ subspaces, which may indicate anomalous data  points. The principal components of PCA are the eigenvectors of the  covariance matrix, which is computed twice to improve robustness  Robust Principal Component Analysis (rPCA)  is based on the Principal Component Analysis (PCA), that is used for dimensionality reduction. PCA is used to detect subspaces in a dataset and has been applied to anomaly detection to identify deviations from the ’expected’ subspaces, which may indicate anomalous data points. The principal components of PCA are the eigenvectors of the covariance matrix, which is computed twice to improve robustness | The low dimensional subspace computed by RPCA, even on training data, is more representative of the true nominal state of the measured data. This allows for a great range of anomalies, and hence network attacks, to be successfully detected. | Its F1 score and Accuracy is less than other techniques, and is slower in larger data sets |
| 1. CMGOS | Goldstein M. Anomaly Detection in Large Datasets [PhD-Thesis]. University of Kaiserslautern. München, Germany; 2014 | In CMGOS, the local density estimation is performed by estimating a multivariate Gaussian model, whereas the Mahalanobis distance serves as a basis for computing the anomaly score.Finally, the CMGOS score is computed by dividing the Mahalanobis distance of an instance to its nearest cluster center by the chi-squared distribution with a certain confidence interval | It can be confirmed that the clustering-based algorithms (except for CMGOS-MCD) are faster than the nearest-neighbor based algorithms with the quadratic search complexity. | Time required for recursive runs is more. |