Anomaly Detection and Application to Intrusion Detection

Introduction

- We are drowning in the deluge of data that are being collected world-wide, while starving for knowledge at the same time*
- Anomalous events occur relatively infrequently
- However, when they do occur, their consequences can be quite dramatic and quite often in a negative sense

What are Anomalies?

- Anomaly is a pattern in the data that does not conform to the expected behavior
- Also referred to as outliers, exceptions, peculiarities, surprises, etc.
- Anomalies translate to significant (often critical) real life entities
  - Cyber intrusions
  - Credit card fraud
  - Faults in mechanical systems

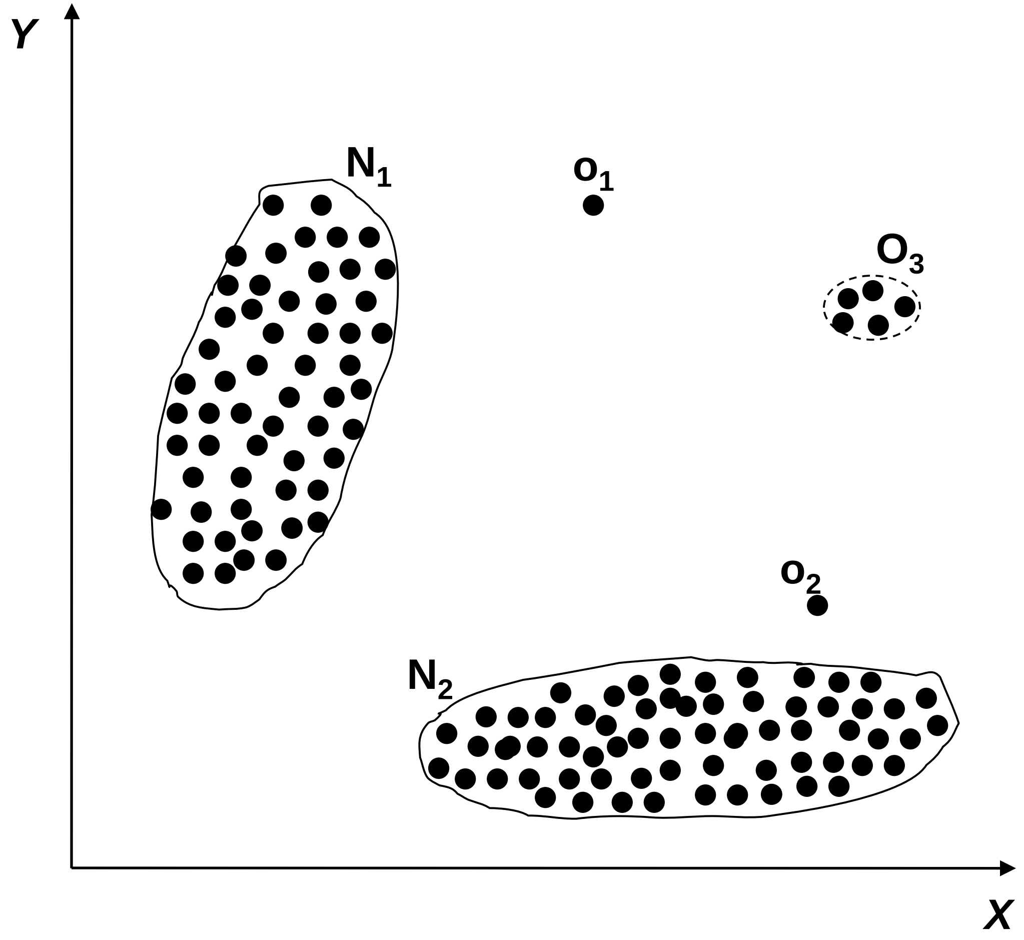
Real World Anomalies

- Credit Card Fraud
  - An abnormally high purchase made on a credit card

- Cyber Intrusions
  - A web server involved in *ftp* traffic

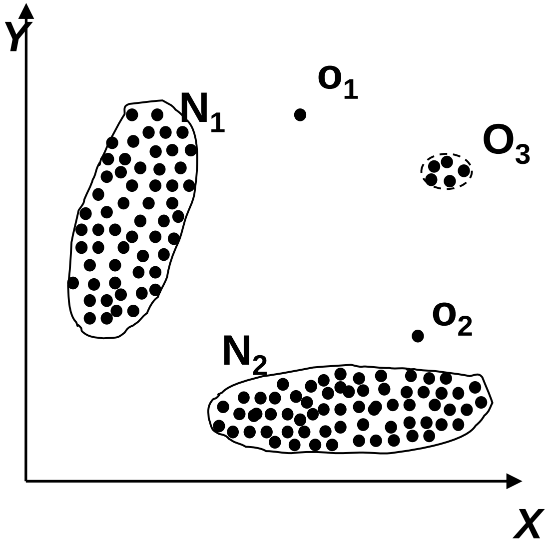
Simple Examples

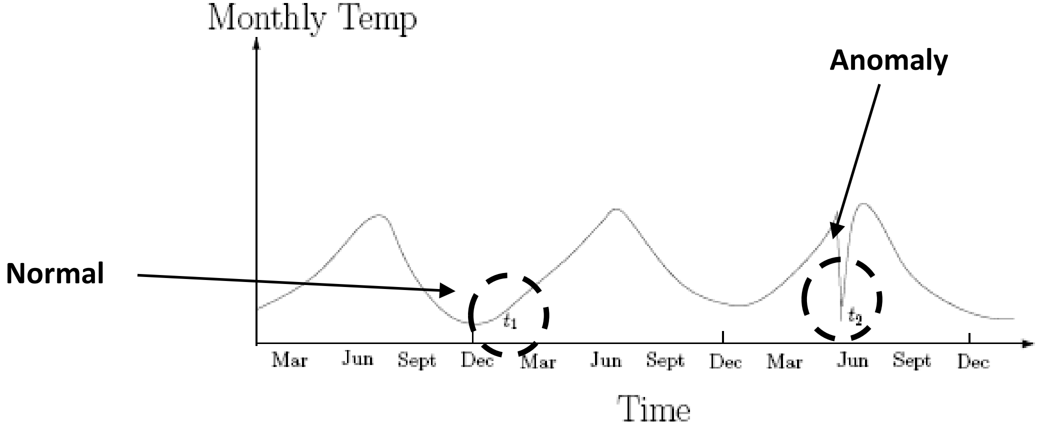
- N_1 and N_2 are regions of normal behavior
- Points o_1 and o_2 are anomalies
- Points in region O_3 are also anomalies

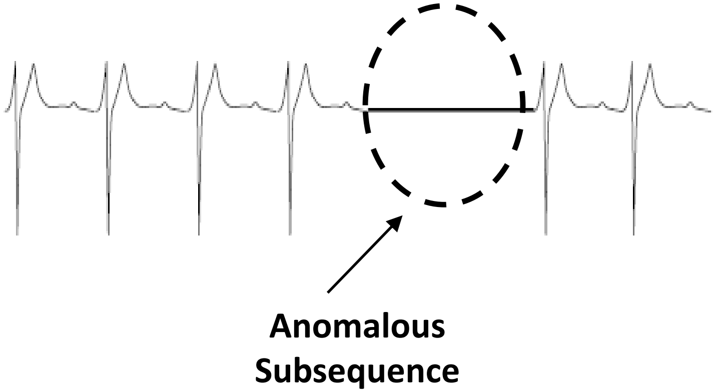


Type of Anomalies

- **Point Anomalies**: An individual data instance is anomalous w.r.t. the data
- **Contextual Anomalies**: An individual data instance is anomalous within a context. Requires a notion of context Also referred to as conditional anomalies
- **Collective Anomalies:** A collection of related data instances is anomalous The individual instances within a collective anomaly are not anomalous by themselves. Requires a relationship among data instances
  - Sequential Data
  - Spatial Data
  - Graph Data







Key Challenges of Anomaly Detection

- **Defining a representative normal region** is challenging
- The **boundary between normal and outlying** behavior is often not precise
- **Availability of labeled data** for training/validation
- The **exact notion of an outlier** is different for different application domains
- Data might contain **noise**
- Normal behavior keeps evolving
- Appropriate **selection of relevant features**

Input Data

- Most common form of data handled by anomaly detection techniques is *Record Data*
  - Univariate
  - Multivariate

Engine Temperature
192
195
180
199
19
177
172
285
195
163

Univariate

Multivariate

Tid	SrcIP	Start time	Dest IP	Dest Port	Number of bytes	Attack
1	206.135.38.95	11:07:20	160.94.179.223	139	192	No
2	206.163.37.95	11:13:56	160.94.179.219	139	195	No
3	206.163.37.95	11:14:29	160.94.179.217	139	180	No
4	206.163.37.95	11:14:30	160.94.179.255	139	199	No
5	206.163.37.95	11:14:32	160.94.179.254	139	19	Yes
6	206.163.37.95	11:14:35	160.94.179.253	139	177	No
7	206.163.37.95	11:14:36	160.94.179.252	139	172	No
8	206.163.37.95	11:14:38	160.94.179.251	139	285	Yes
9	206.163.37.95	11:14:41	160.94.179.250	139	195	No
10	206.163.37.95	11:14:44	160.94.179.249	139	163	Yes

Data Labels

- Supervised Anomaly Detection
  - Labels available for both normal data and anomalies
  - Similar to rare class mining
  - Machine learning models: Naïve Bayes, Neural Network
- Semi-supervised Anomaly Detection
  - Labels available only for normal data
- Unsupervised Anomaly Detection
  - No labels assumed
  - Based on the assumption that anomalies are very rare compared to normal data
  - Machine learning: clustering

Output of Anomaly Detection

- Label
  - Each test instance is given a *normal* or *anomaly* label
  - This is especially true of classification-based approaches
- Score
  - Each test instance is assigned an anomaly score
    - Allows the output to be ranked
    - Requires an additional threshold parameter

Evaluation of Anomaly Detection – F-value

- Accuracy is not sufficient metric for evaluation
  - Example: network traffic data set with 99.9% of normal data and 0.1% of intrusions
  - Trivial classifier that labels everything with the normal class can achieve 99.9% accuracy !!!!!

Confusion matrix		Predicted class	
		NC	C
Actual class	NC	TN	FP
	C	FN	TP

anomaly class – C

normal class – NC

- **Focus on both recall and precision**
  - **Recall (R) = TP/(TP + FN)**
  - **Precision (P) = TP/(TP + FP)**
- **F – measure = 2*R*P/(R+P)**

Applications of Anomaly Detection

- Network intrusion detection
- Insurance / Credit card fraud detection
- Healthcare Informatics / Medical diagnostics
- Industrial Damage Detection
- Image Processing / Video surveillance
- Novel Topic Detection in Text Mining

Intrusion Detection

- Intrusion Detection:
  - Process of monitoring the events occurring in a computer system or network and analyzing them for intrusions
  - Intrusions are defined as attempts to bypass the security mechanisms of a computer or network
- Challenges
  - Traditional signature-based intrusion detection systems are based on signatures of known attacks and cannot detect emerging cyber threats
  - Substantial latency in deployment of newly created signatures across the computer system
- Anomaly detection can alleviate these limitations

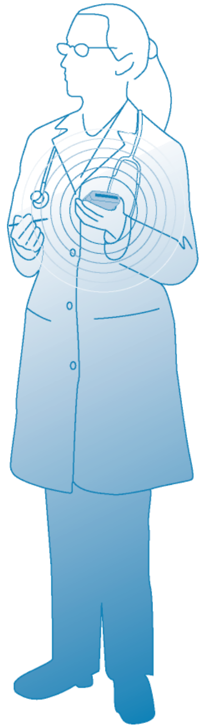
Fraud Detection

- Fraud detection refers to detection of criminal activities occurring in commercial organizations
  - Malicious users might be the actual customers of the organization or might be posing as a customer (also known as identity theft).
- Types of fraud
  - Credit card fraud
  - Insurance claim fraud
  - Mobile / cell phone fraud
  - Insider trading

- Challenges
  - Fast and accurate real-time detection
  - Misclassification cost is very high

Healthcare Informatics

- Detect anomalous patient records
  - Indicate disease outbreaks, instrumentation errors, etc.
- Key Challenges
  - Only normal labels available
  - Misclassification cost is very high
  - Data can be complex: spatio-temporal



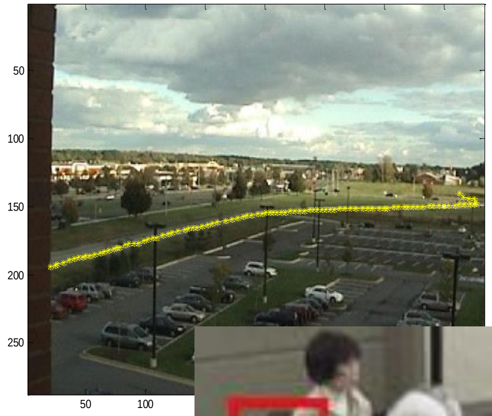
Industrial Damage Detection

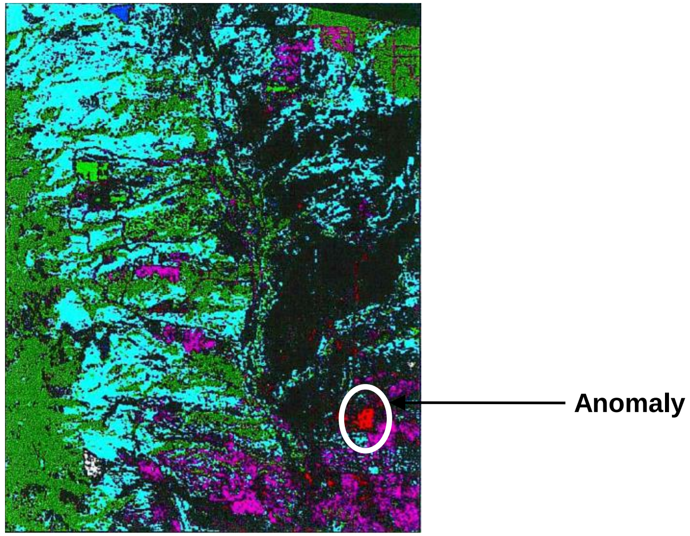
- Industrial damage detection refers to detection of different faults and failures in complex industrial systems, structural damages, intrusions in electronic security systems, abnormal energy consumption, etc.
  - Example: Aircraft Safety
    - Anomalous Aircraft (Engine) / Fleet Usage
    - Anomalies in engine combustion data
    - Total aircraft health and usage management

- Key Challenges
  - Data is extremely huge, noisy and unlabelled
  - Most of applications exhibit temporal behavior
  - Detecting anomalous events typically require immediate intervention

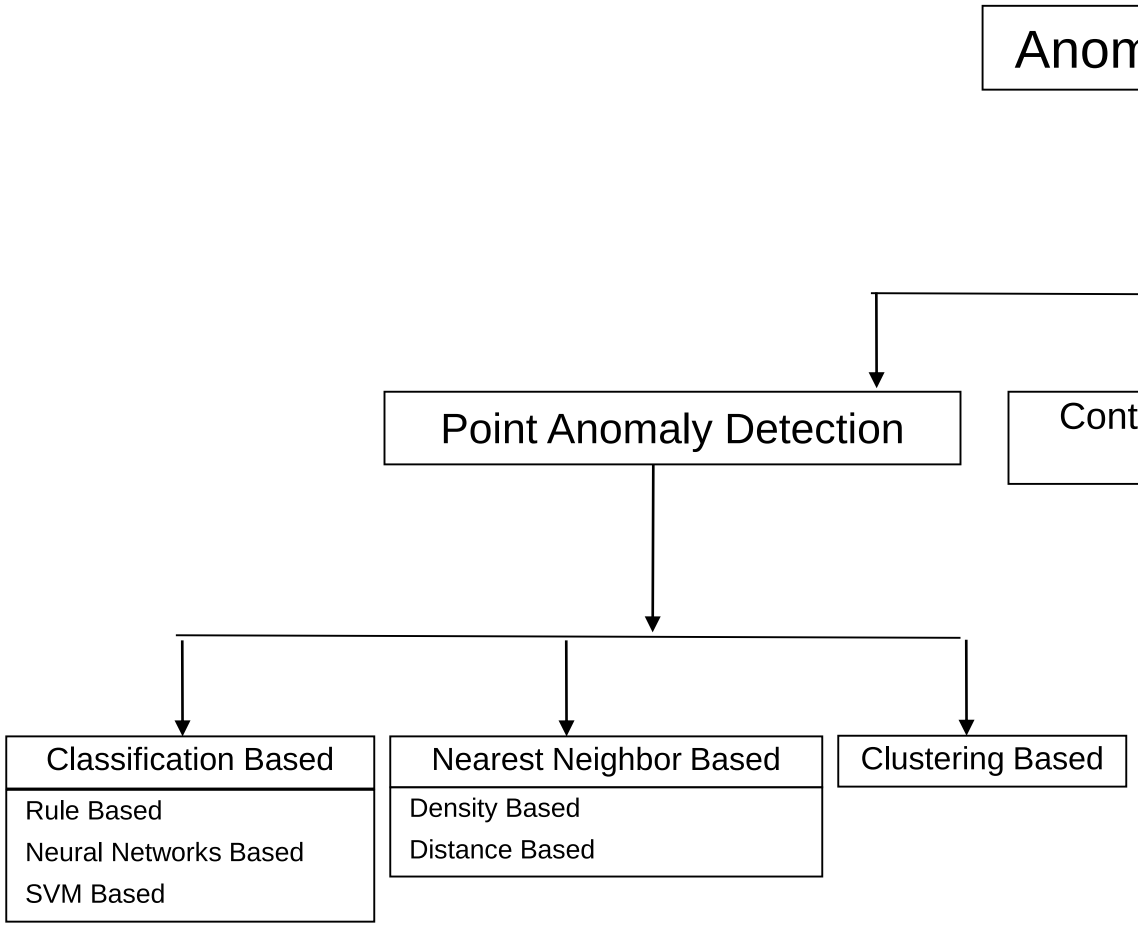
Computer Vision

- Detecting outliers in a image or video monitored over time
- Detecting anomalous regions within an image
- Used in
  - mammography image analysis
  - video surveillance
  - satellite image analysis
- Key Challenges
  - Detecting collective anomalies
  - Data sets are very large





Taxonomy


Anomaly Detection
Point Anomaly Detection
Contextual Anomaly Detection
Collective Anomaly Detection
Classification Based
Rule Based
Neural Networks Based
SVM Based
Nearest Neighbor Based
Density Based
Distance Based
Clustering Based

Classification Based Techniques

- Main idea: build a classification model for normal (and anomalous (rare)) events based on labeled training data, and use it to classify each new unseen event
- Classification models must be able to handle skewed (imbalanced) class distributions
- Categories:
  - *Supervised classification techniques*
    - Require knowledge of both ***normal*** and ***anomaly*** class
    - Build classifier to distinguish between normal and known anomalies

Nearest Neighbor Based Techniques

- *Key assumption*: normal points have close neighbors while anomalies are located far from other points
- General two-step approach
  1. Compute neighborhood for each data record
  2. Analyze the neighborhood to determine whether data record is anomaly or not
- Categories:
  - Distance based methods
    - Anomalies are data points most distant from other points
  - Density based methods
    - Anomalies are data points in low density regions

Nearest Neighbor Based Techniques

- Advantage
  - Can be used in unsupervised or semi-supervised setting (do not make any assumptions about data distribution)
- Drawbacks
  - If normal points do not have sufficient number of neighbors the techniques may fail
  - Computationally expensive
  - In high dimensional spaces, data is sparse and the concept of similarity may not be meaningful anymore. Due to the sparseness, distances between any two data records may become quite similar => Each data record may be considered as potential outlier!

Nearest Neighbor Based Techniques

- Distance based approaches
  - A point *O* in a dataset is an *DB*(*p*, *d*) outlier if at least fraction *p* of the points in the data set lies greater than distance *d* from the point *O**
- Density based approaches
  - Compute local densities of particular regions and declare instances in low density regions as potential anomalies
  - Approaches
    - Local Outlier Factor (LOF)
    - Connectivity Outlier Factor (COF)
    - Multi-Granularity Deviation Factor (MDEF)

Clustering Based Techniques

- *Key Assumption:* Normal data instances belong to large and dense clusters, while anomalies do not belong to any significant cluster.
- *General Approach:*
  - Cluster data into a finite number of clusters.
  - Analyze each data instance with respect to its closest cluster.
  - Anomalous Instances
    - Data instances that **do not fit** into any cluster (residuals from clustering).
    - Data instances in **small clusters**.
    - Data instances in **low density clusters**.
    - Data instances that are **far from other points within the same cluster**.

Clustering Based Techniques

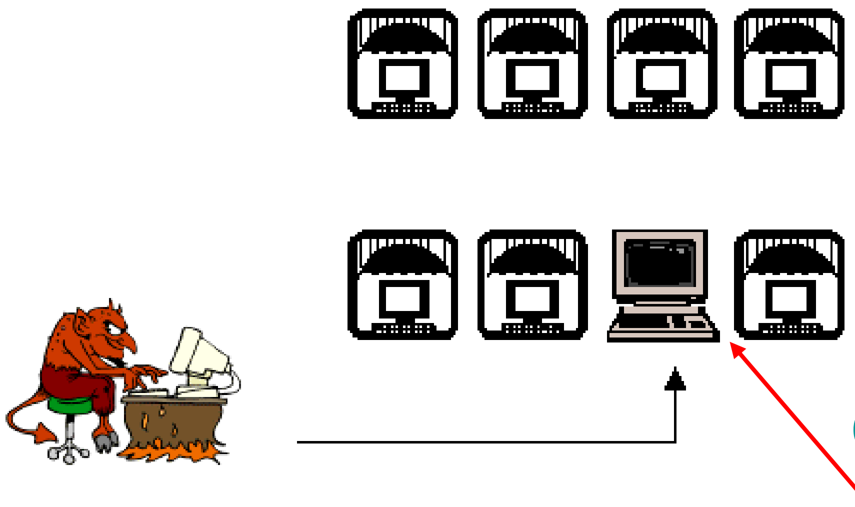
- Advantages
  - Unsupervised algorithm
  - Existing clustering algorithms can be plugged in
- Drawbacks
  - If the data does not have a natural clustering or the clustering algorithm is not able to detect the natural clusters, the techniques may fail
  - Computationally expensive
  - In high dimensional spaces, data is sparse and distances between any two data records may become quite similar

What are Intrusions?

- Intrusions are actions that attempt to bypass security mechanisms of computer systems. They are usually caused by:
  - Attackers accessing the system from Internet
  - Insider attackers - authorized users attempting to gain and misuse non-authorized privileges
- Typical intrusion scenario

Scanning activity

Computer Network



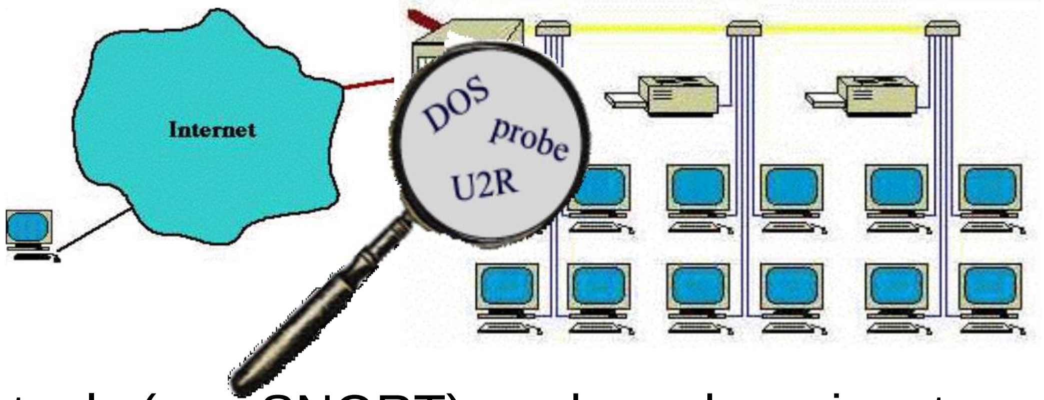
Compromised Machine

Machine with vulnerability

Attacker

Intrusion Detection

- Intrusion Detection System
  - combination of software and hardware that attempts to perform intrusion detection
  - raises the alarm when possible intrusion happens



- Traditional intrusion detection system IDS tools (e.g. SNORT) are based on signatures of known attacks
  - Example of SNORT rule (MS-SQL “Slammer” worm)

    any -> udp port 1434 (content:"|81 F1 03 01 04 9B 81 F1 01|"; content:"sock"; content:"send")

SNORT

www.snort.org

- Limitations
  - Signature database has to be manually revised for each new type of discovered intrusion
  - They cannot detect emerging cyber threats
  - Substantial latency in deployment of newly created signatures across the computer system

Approaches for Intrusion Detection

- Increased interest in machine learning based intrusion detection
  - Attacks for which it is difficult to build signatures
  - Attack stealthiness
  - Unforeseen/Unknown/Emerging attacks
  - Distributed/coordinated attacks
- Approaches for intrusion detection
  - *Misuse detection (Classification)*
    - Building predictive models from labeled data sets (instances are labeled as “normal” or “intrusive”) to identify known intrusions
    - High accuracy in detecting many kinds of known attacks
    - Cannot detect unknown and emerging attacks
  - *Apply Anomaly detection*
    - Detect novel attacks as deviations from “normal” behavior
    - Potential high false alarm rate - previously unseen (yet legitimate) system behaviors may also be recognized as anomalies
  - *Summarization of network traffic (Rule-based)*

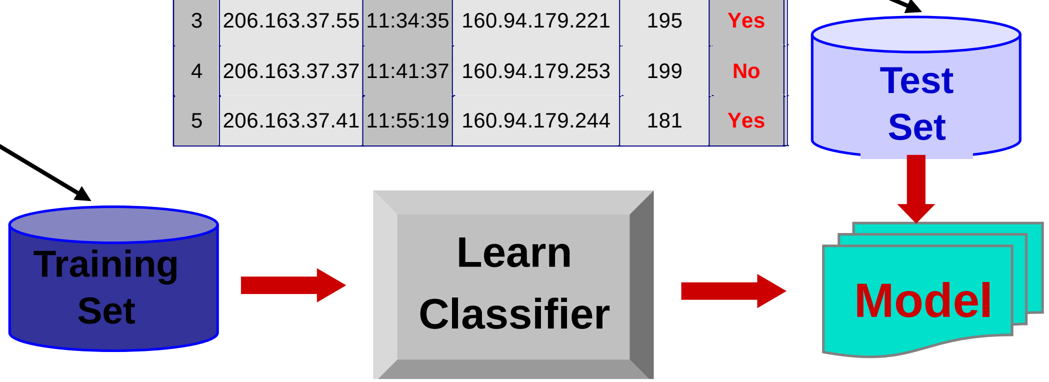
Approaches for Intrusion Detection

Misuse Detection – Building Predictive Models

Tid	SrcIP	Start time	Dest IP	Dest Port	Number of bytes	Attack
1	206.135.38.95	11:07:20	160.94.179.223	139	192	No
2	206.163.37.95	11:13:56	160.94.179.219	139	195	No
3	206.163.37.95	11:14:29	160.94.179.217	139	180	No
4	206.163.37.95	11:14:30	160.94.179.255	139	199	No
5	206.163.37.95	11:14:32	160.94.179.254	139	19	Yes
6	206.163.37.95	11:14:35	160.94.179.253	139	177	No
7	206.163.37.95	11:14:36	160.94.179.252	139	172	No
8	206.163.37.95	11:14:38	160.94.179.251	139	285	Yes
9	206.163.37.95	11:14:41	160.94.179.250	139	195	No
10	206.163.37.95	11:14:44	160.94.179.249	139	163	Yes

categorical temporal categorical continuous class

Tid	SrcIP	Start time	Dest IP	Number of bytes	Attack
1	206.163.37.81	11:17:51	160.94.179.208	150	No
2	206.163.37.99	11:18:10	160.94.179.235	208	No
3	206.163.37.55	11:34:35	160.94.179.221	195	Yes
4	206.163.37.37	11:41:37	160.94.179.253	199	No
5	206.163.37.41	11:55:19	160.94.179.244	181	Yes

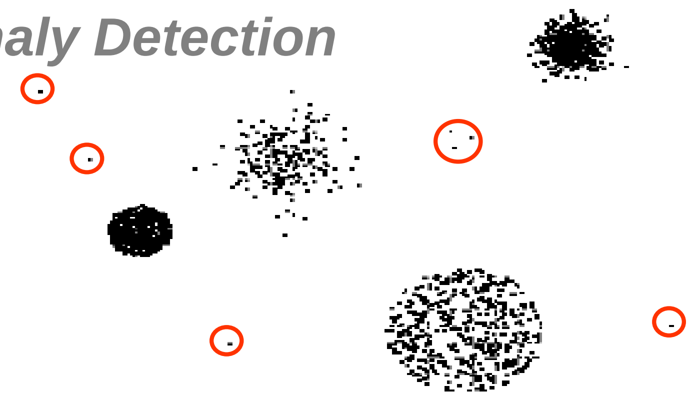


Summarization of attacks using association rules

Rules Discovered:

{Src IP = 206.163.37.95,
Dest Port = 139,
Bytes ∈ [150, 200]} --> {ATTACK}

Anomaly Detection



Feature Extraction

- Three groups of features
  - Basic features of individual TCP connections
    - source & destination IP
    - source & destination port
    - Protocol
    - Duration
    - Bytes per packets
    - number of bytes
  - **Time based features**
    - For the same source (destination) IP address, number of unique destination (source) IP addresses inside the network *in last T seconds –*
    - Number of connections from source (destination) IP to the same destination (source) port *in last T seconds*
  - Connection based features
    - For the same source (destination) IP address, number of unique destination (source) IP addresses inside the network *in last N connections*
    - Number of connections from source (destination) IP to the same destination (source) port *in last N connections*

Conclusions

- Anomaly detection can detect critical information in data.
- Highly applicable in various application domains.
- Nature of anomaly detection problem is dependent on the application domain.
- Need different approaches to solve a particular problem formulation.