What is Anomaly Detection?

An anomaly, in cybersecurity terms, refers to any data, entity, observation, or behavior that deviates from the norm or shows an unexpected change in a dataset.

Anomaly detection (or outlier detection) has been used in the statistics industry for years and is also a part of human and animal behavior, such as detecting ripe fruit or rotten vegetables.

In businesses, anomalous behavior or data can be positive but usually represents a threat.

An anomaly detection example: a spike in sales after optimizing an e-commerce app is a positive anomaly, while too many transactions from a bank account within a short period could signal a threat.

Anomalies can occur in two ways:

- **Unexpected behavior**, such as a sudden spike in website traffic without any specific reason.
- **Absence of expected behavior**, like an e-commerce site not achieving its usual sales on a peak day, such as Cyber Monday.

The context of an anomaly varies across businesses, depending on their standard metrics or typical data patterns. Anomalies are not inherently ‘good’ or ‘bad’; they are just deviations from the usual or the expected.

Businesses must differentiate between a genuine anomaly, which indicates concern or opportunity, such as hacking, equipment malfunctions, or sales spikes, and a **false positive**, such as irrelevant changes or noise that can be ignored.

Importance of Anomaly Detection in your Business

Anomaly detection is crucial for ensuring smooth business operations, tracking performance, and maintaining data and system security. Detecting these unusual patterns at the earliest will help businesses address issues before they escalate.

1. Enhancing Business Operations and Decision-Making

Anomaly detection helps identify areas for improvement, threats, and growth opportunities. It can alert organizations to equipment failures, pricing issues, or fraudulent activities. Real-time detection of KPIs, such as sales spikes, allows businesses to react quickly and optimize operations.

2. Preventing Fraud and Security Threats

Anomaly detection systems can quickly spot unusual data or behaviors, like hacking, fraud, or security threats. By tracking odd login patterns or traffic spikes, they give early warnings to help reduce risks and prevent breaches.

3. Optimizing IT and Application Performance

Anomaly detection helps maintain high performance in IT systems and applications by identifying slow response times or system overloads. This proactive approach prevents disruptions and ensures smooth business operations.

4. Improving Product Quality

By monitoring product performance or customer experience, anomaly detection quickly identifies issues like malfunctions or unexpected behavior. It helps resolve problems promptly, protecting the brand’s reputation and revenue.

5. Cloud Cost Management

Anomaly detection helps manage cloud costs by identifying unexpected cost spikes. By analyzing past data, it alerts the right people to inefficiencies, helping them optimize resources and reduce costs.

4 Keys to Automating Threat Detection, Threat Hunting and Response

- Maturing Advanced Threat Defense
- 4 Must-Do’s for Advanced Threat Defense
- Automating Detection and Response

Download the Whitepaper Now!

Anomaly Detection in Time Series Data

Time series anomaly detection is a method used to point out unusual or unexpected data points or patterns in data collected over time. Each data point includes a timestamp and its associated value. And anomaly detection systems use this normal behavior to spot unusual events and give alerts on key issues in key performance indicators (KPIs).

Time series anomaly detection is useful for tracking important business metrics over time, such as mobile app installs, web page views, cost per click, and bounce rate. These detection systems establish a baseline of normalcy for key metrics and monitor data for seasonal or cyclical patterns. The ability to automate this process is essential when handling large datasets across multiple metrics, enabling businesses to detect anomalies efficiently and uncover valuable

insights.

Benefits of Time Series Anomaly Detection

- **Provides real-time alerts:** This method can quickly identify abnormal trends in KPIs and raise issues even before they escalate, helping businesses take immediate action.
- **Higher accuracy:** It can detect even subtle changes in patterns in time-based data points, helping businesses with actionable insights.
- **Automated anomaly detection:** This detection method can reduce manual monitoring efforts by automating detection, even across larger datasets.
- **Scalability:** As businesses grow, a time series anomaly based detection system can scale accordingly to handle large sets of data.

Different Types of Time Series Anomalies

Anomalies in a business can fall under any of the categories below.

Global outliers/point anomalies

Global outliers are data points that are far outside the usual patterns, either accidentally or on purpose.

For example, a customer transfers a relatively large amount of money from his account, which is unusual compared to all his previous transactions so far.

Contextual outliers

These are data points that deviate from the usual pattern within a specific context, even though this deviation is normal when considered individually.

For example, a customer who normally does online shopping during the day suddenly makes a bulk purchase at an odd time, like 3 a.m., which is unusual. This could be an anomaly when considering the time of day and the customer’s usual shopping behavior in that shopping app.

Collective Outliers

Collective outliers occur when a group of data points deviate from the norm when viewed together.

For instance, several customers who typically make small purchases may suddenly buy large quantities of the same product within a short time frame. This could indicate a flash sale or even fraudulent activity like a coordinated resale effort.

Anomaly Detection Techniques

There are mainly 3 anomaly detection techniques:

1. Visualization

Visualization methods involve creating charts, graphs, or plots to make data patterns easier to spot. Data analysts can then visually inspect the data for any points that differ from the expected or normal patterns, identifying them as anomalies. This method is useful for initial

analysis.

2. Statistical Tests

Statistical tests detect anomalies by comparing the actual data against expected patterns or distributions. These tests help identify when specific data points are significantly different from the norm. Common statistical data anomaly detection tests include Z-tests, T-tests, and Chi-square tests. They are often used when data is assumed to follow a known distribution (e.g., normal distribution)

3. Machine Learning Algorithms

Machine learning algorithms detect anomalies by learning the underlying patterns in the data and identifying deviations from these patterns. Some common **machine-learning** techniques for anomaly detection include:

- **Decision Trees (Isolation Forest):** This ensemble method isolates anomalies by partitioning the data. It works by randomly selecting a feature and splitting it into subsets, with anomalies being isolated quickly compared to normal points. This technique is particularly effective for high-dimensional data.
- **One-Class SVM:** This method is trained on 'normal' data and creates a boundary around it. Anything outside this boundary is considered an anomaly. It's particularly useful when the dataset has few or no labelled anomalies.
- **K-Means Clustering:** K-Means groups data points into clusters based on their proximity to a centroid. Points that are far from any cluster centroid are considered anomalies, as they don't fit into any cluster

What Are the Types of Anomaly Detection in Machine Learning?

There are 3 types of machine learning-based anomaly detection:

1. Supervised Anomaly Detection Technique

This anomaly detection model uses an algorithm trained on a labelled dataset that includes both normal and anomalous data. These techniques are rarely used because labelled data is hard to obtain, and the data typically has an imbalance, with many more normal instances than anomalies.

Common supervised methods include Bayesian networks, k-nearest neighbors, decision trees, supervised neural networks, and SVMs.

Supervised models may offer a higher detection rate because they can return a confidence score and incorporate prior knowledge and interdependencies between variables.

2. Semi-Supervised Anomaly Detection Technique

This technique uses a normally labelled training dataset to construct a model representing normal behavior. The model is then used to detect anomalies by testing how likely the model is to generate any new instance.

“Semi-supervised” also describes a method where a dataset has some labelled data. The model uses this labelled portion to create a classification algorithm and then predicts the labels for the unlabelled data.

3. Unsupervised Anomaly Detection Techniques

This type of anomaly detection finds unusual patterns in data that don’t have labels, using the data’s own characteristics. These methods are widely used because they can work in many situations, but they need a lot of data and computing power.

Popular unsupervised **anomaly detection algorithms** include Autoencoders, K-means, GMMs, hypothesis tests-based analysis, and PCAs.

Compared to supervised anomaly detection, unsupervised anomaly detection works best for businesses needing real-time monitoring and quick responses while dealing with large datasets. Traditional anomaly detection methods might miss these sudden activity jumps, but techniques like cluster analysis can identify them more easily.

Why Should You Use Anomaly Detection

- **Fraud Detection**

Used in banking, insurance, and trading to identify unauthorized transactions, money laundering, and abnormal trading patterns in real-time.

- **Cybersecurity & Network Security**

Anomaly detection in cyber security and network security is used to identify suspicious and unusual network traffic patterns, helping to detect security threats such as malware or unauthorized access. This is achieved through Intrusion Detection Systems (IDS) and Network Detection and Response (NDR) solutions, like Fidelis Network®.

- **Manufacturing & Quality Control**

In manufacturing, anomaly detection monitoring, paired with computer vision, helps spot defects or packaging issues by analyzing sensor data, camera footage, and production metrics.

- **IT Systems Management**

Anomaly detection monitors IT system performance by identifying unusual patterns in server logs, helping predict failures, and ensuring smooth operations.

- **Energy, Transportation & Critical Infrastructure**

Predicts equipment failures and optimizes maintenance by monitoring data from IoT sensors and operational technology devices.

- **Retail & E-commerce**

Allows merchants to spot threats like fraud, fake reviews, and irregular purchasing patterns. Beyond identifying these risks, it also helps predict customer churn and optimize marketing strategies.

In network security, anomaly detection can be improved with advanced tools like Fidelis Network

® Detection and Response (NDR), offering real-time insights.

How Fidelis Network® NDR Enhances Anomaly Detection

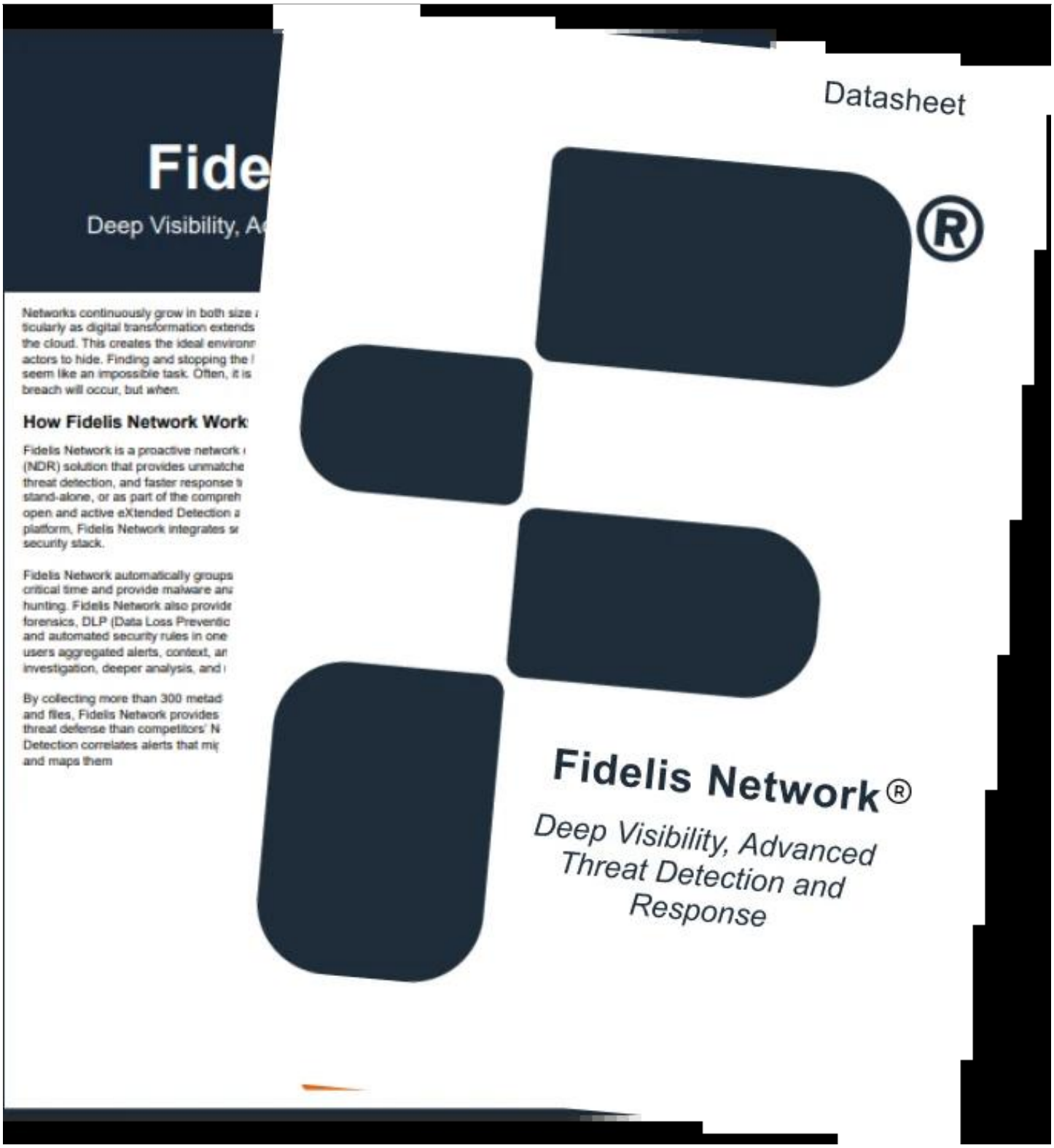
Fidelis Network® Detection and Response (NDR) provides complete visibility into your business's network activities by thoroughly monitoring all ports and protocols. Its advanced techniques detect abnormal patterns that could point to security threats, such as unauthorized access or other malicious actions. By continuously tracking normal network behavior, Fidelis NDR can spot changes that may signal emerging risks, enabling quick action to reduce the impact of security breaches.

Through its patented traffic analysis tools and risk-aware terrain mapping, Fidelis NDR provides organizations with powerful, real-time capabilities for detecting anomalous behavior and proactively addressing network security challenges.

Transform your security with Fidelis NDR

- Proactive Network Data Loss Prevention
- Threat analysis with precision Sandboxing
- Comprehensive Cyber Terrain Mapping

Download Datasheet



Frequently Ask Questions

What is anomaly detection and why is it important for businesses?

Anomaly detection means finding data or actions that are different from normal. It helps businesses spot problems like fraud or equipment breakdowns early, so they can fix them before things get worse.

What are the main techniques used to detect anomalies?

- **Visualization:** Using charts to manually spot irregular patterns.
- **Statistical Tests:** Comparing data to expected patterns using tools like Z-tests.
- **Machine Learning Algorithms:** Training models to automatically find patterns and detect unusual data.

What are the types of anomaly detection in machine learning?

- **Supervised:** Uses labeled data to detect known anomalies.
- **Semi-Supervised:** Trains on normal data and flags anything that looks unusual.
- **Unsupervised:** Doesn’t need labeled data and finds anomalies based on data patterns alone.

How does anomaly detection help in cybersecurity and IT systems?

It monitors network traffic and system logs to spot suspicious activity, like unauthorized access or malware. This helps prevent cyberattacks and keeps IT systems running smoothly.