Veeam Ransomware Detection Service

Payam Kavousi

Senior Software Engineer, Veeam

Zack Rossman

Staff Software Engineer, Veeam

Abstract

This document outlines the technical foundation of Veeam's Ransomware Detection Service, detailing the methodologies employed for detecting ransomware attacks via Random Cut Forest machine learning models.

Intended Audience

This white paper is designed for professionals responsible for planning, defending against, and remediating ransomware attacks, including:

- Chief information officers (CIOs)
- Chief information security officers (CISOs)
- IT managers
- Cybersecurity specialists
- Backup administrators

Executive Summary

The Challenge

With 69% of organizations experiencing at least one ransomware incident in 2024 and 2025 (2025 Veeam Ransomware Trend and Proactive Strategies survey), traditional backup solutions alone are no longer sufficient. Organizations need intelligent detection that can distinguish between legitimate business operations and malicious encryption before it's too late.

The Solution

Veeam's Ransomware Detection Service uses machine learning to automatically analyze backup data and identify ransomware activity with high accuracy. The system protects against both rapid attacks and sophisticated, slow-moving threats through multiple detection profiles.

Key Benefits

Feature	Value
Personalized Threat Detection	Individual models per user/resource recognize that each has unique patterns. This granular approach catches anomalies that global models would miss.
Privacy-Preserving Architecture	Anomaly detection operates solely on derived statistics (e.g., file counts, entropy scores). File content stays within the backup system and never enters the ML model.
Comprehensive Coverage	Detects multiple strains of ransomware attacks (e.g. fast "blitz" attacks and slow-moving "stealth" attacks).
Real-Time Unsupervised Learning	Unsupervised learning incrementally adapts to your environment in real-time. No need for manual configuration or re-train models from scratch.

Business Value

Early ransomware detection minimizes business disruption and accelerates recovery by identifying threats before they can spread through your backup infrastructure. The system reduces IT security workload through intelligent automation while also helping organizations meet compliance requirements for threat detection and response. Continuous adaptation protects against evolving ransomware tactics without requiring manual updates, making it easier to safeguard your critical data assets as threats change.

Introduction

Understanding Ransomware

Ransomware is a form of cyberattack where threat actors encrypt organizational data and demand payment for its restoration. These attacks disrupt or suspend operations and therefore force management to choose between paying the ransom or attempting independent recovery.

Current Threat Landscape

The 2025 Veeam Ransomware Trend and Proactive Strategies survey, encompassing more than 1,300 organizations worldwide, revealed that 69% of those organizations were impacted by at least one ransomware incident that resulted in encryption or data exfiltration.

Although significant ransomware groups such as LockBit, BlackCat, and Black Basta have been dismantled by law enforcement, there has still been an increase in smaller groups and lone-wolf threat actors targeting small and medium-sized enterprises (SMEs) that often lack robust cybersecurity defenses.

Veeam's Dynamic Attack Profile Approach

Veeam uses AI to detect ransomware attacks by looking for unusual changes in your backups.

The system monitors backup files for suspicious signals like:

- Anomalous file modifications or deletion.
- Suspect file extensions.
- Unusual patterns in file data.

It works automatically without needing pre-programmed rules and learns what's "normal" for your environment and alerts you when something looks wrong. This helps customers quickly spot ransomware attacks and restore clean data.

The technology behind it is called Random Cut Forest (RCF), an Amazon-developed algorithm that's designed to find anomalies in real-time data streams.

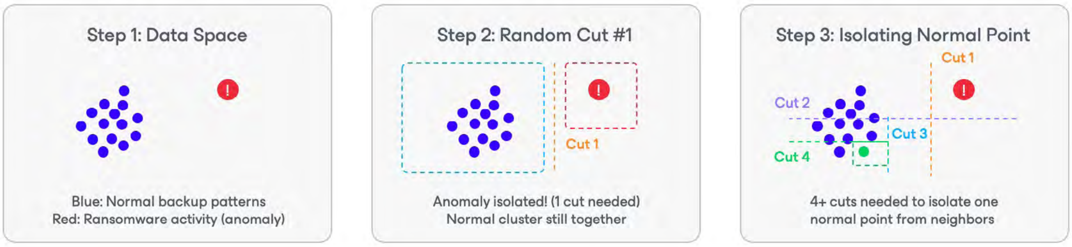
Detection Methodology

Random Cut Forest Algorithm

Veeam's ransomware detection relies on the RCF machine learning method, which is an unsupervised anomaly detection algorithm that's specifically optimized for streaming data analysis. RCF requires no labeled training data and continuously learns normal behavior patterns from streaming backup data. It's worth noting that each personalized threat detection model is tailored to each protected object/entity/user and this improves accuracy.

Random Cut Forest: How Anomalies Are Detected

Anomalies require fewer random cuts to isolate than normal points


Step 1: Data Space
!
Blue: Normal backup patterns
Red: Ransomware activity (anomaly)
Step 2: Random Cut #1
!
Cut 1
Anomaly isolated! (1 cut needed)
Normal cluster still together
Step 3: Isolating Normal Point
Cut 1
!
Cut 2
Cut 3
Cut 4
4+ cuts needed to isolate one
normal point from neighbors

RCF Algorithm: Key Principle

How Random Cut Forest Works:

1. Randomly select a dimension (feature) and value
2. Split the data space at that point
3. Repeat until each point is isolated
4. Count cuts needed for each point

→ Fewer cuts = Higher anomaly score

Applied to Veeam Backups:

- Dimensions: File entropy, extensions, modifications
- Normal backups: Form dense clusters
- Ransomware: Creates outlier patterns
- Result: Ransomware isolated quickly

→ Automatic detection without training data

Why This Works for Ransomware Detection

Ransomware creates unusual patterns (high entropy, strange extensions) that stand out from normal file changes.

Dynamic Attack Detection

Veeam's Ransomware Detection Service moinitors backup data by using two complementary detection approaches to identify different attack patterns.

Detection Profiles

Profile	Purpose
Fast Attack	Detects rapid encryption events
Extended Attack	Identifies gradual, persistent threats

Fast Attack Profile

Identifies sudden, high-volume encryption activity. This catches aggressive attacks that aim for immediate impact.

Extended Attack Profile

This monitors patterns over consecutive backups to detect slower, more sophisticated threats that attempt to avoid detection through gradual activity.

Both profiles run simultaneously to provide comprehensive coverage against various attack methodologies. This is a best-practice known as "ensemble methods" defined as machine learning techniques that combine predictions from multiple individual models (often called "weak learners") to create a single, more accurate, and robust model.

Feature Analysis

The detection service analyzes backup metadata to identify potential threats by monitoring file characteristics and modification patterns.

Detection Signals

Through research and testing, we identified signals related to the entropy of a file's bytes and other metadata which can reliably differentiate between files that have been ransomware encrypted vs. non-ransomware encrypted. These signals are continuously analyzed across backup cycles to detect anomalous patterns.

Threat Reporting

When multiple detection signals identify the same potential threat, the system consolidates findings to provide a single, prioritized alert. This approach ensures clear reporting and minimizes redundant notifications.

The service continuously monitors backup data and alerts when file changes deviate significantly from normal patterns.

In summary, Veeam's ransomware detection service maintains multiple models for each protected object/entity/user. Normally, this would have created cost and scale concerns but online learning and Veeam's patent-pending architecture address these concerns.

Threat Assessment and Management

Threat Status Categories

The system tracks ransomware incidents with the following status types:

Status	Description
in review	Threat is being investigated by Veeam operators (initial state)
active	Threat confirmed and requires immediate attention
resolved	Threat has been addressed or mitigated
false positive	Determined to be a false alarm

Human-in-the-Loop Verification

When potential ransomware activity is detected, the system initially marks threats as in_review. Security operators then analyze these incidents to verify whether they represent genuine threats or false positives before escalating to active status.

This verification step helps ensure accurate threat detection while minimizing false alarms that could impact legitimate business operations.

Status Management

Threat status can be updated through the API as investigations progress and incidents are resolved or confirmed as false positives. The API is secured with strong role-based access control (RBAC) to ensure only authorized personnel can modify threat statuses.

Conclusion

Veeam's Ransomware Detection Service delivers proactive protection against ransomware through intelligent machine learning that adapts to your environment. By combining dual-profile detection with human verification, the system is able to identify likely ransomware threats in an accurate, real-time, and cost-efficient manner.

Key advantages include:

- **Automated intelligence:** Continuously learns normal patterns without manual configuration.
- **Comprehensive coverage:** Protects against both aggressive and stealthy attack methods.
- **Reduced response time:** Enables faster threat identification and clean data restoration.
- **Future-proof protection:** Adapts to emerging ransomware variants without requiring updates.

Organizations can gain peace of mind in knowing their backup infrastructure includes an intelligent layer of defense that works alongside existing security measures to detect threats before they cause irreversible damage.

About Veeam Software

Veeam®, the #1 global market leader in data resilience, believes every business should be able to bounce forward after a disruption with the confidence and control of all their data whenever and wherever they need it. Veeam calls this radical resilience, and we're obsessed with creating innovative ways to help our customers achieve it. Veeam solutions are purpose-built for powering data resilience by providing data backup, data recovery, data portability, data security, and data intelligence. With Veeam, IT and security leaders rest easy knowing that their apps and data are protected and always available across their cloud, virtual, physical, SaaS, and Kubernetes environments. Headquartered in Seattle with offices in more than 30 countries, Veeam protects over 550,000 customers worldwide, including 82% of the Fortune 500, that trust Veeam to keep their businesses running. Radical resilience starts with Veeam. Learn more at www.veeam.com or follow Veeam on LinkedIn @veeam-software and X @veeam.

→ Learn more: veeam.com