INSTITUTE OF INFORMATION
AND COMMUNICATION TECHNOLOGIES
BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **26**, No 2
Sofia • 2026 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2026-0021

A Semantic-Aware File Metadata Generation Framework for Disk-Level Anomaly Detection in Virtual Machine Backups

Jyoti Metan[1], *Mahantesh Mathapati*[2], *Aishwarya Madhusudan*[3], *Santhosh Kumar Gorva*[4], *Bharath Basavaraj*[5], *Benaka Santhosha Siddaiah*[5], *Yogesh Kumaran Selvaraj*[6]

[1]*Department of Information Science & Engineering, Atria Institute of Technology, Bangalore, India*

[2]*Department of Computer Science & Engineering, Amruta Institute of Engineering & Management Sciences, Bangalore, India*

[3]*Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India*

[4]*Department of CSE (Data science), Dayananda Sagar University, Bangalore, India*

[5]*Department of Computer Science and Engineering, Dayananda Sagar University, Bangalore, India*

[6]*Department of Computer Science and Engineering, JAIN (Deemed-to-be University), Bengaluru, India*

***E-mails*:** *jyoti.m@atria.edu mahantesh@aiems.edu.in aishwarya.m@vvce.ac.in santhoshg-ds@dsu.edu.in bharathonly.1992@gmail.com benaka.santhosh-cse@dsu.edu.in yogesh.ks@jainuniversity.ac.in*

***Abstract*:** *The fast growth of Virtual Machine (VM) backup systems in cloud and enterprise environments has greatly led to exposures to disk-level anomalies brought about by ransomware and malicious data corruption. The currently used anomaly detectors are mainly content-based scanning or coarse-grained metadata analysis, which causes high computational complexity, slow response time, and an inability to scale to large-scale backup settings. To combat the above difficulties, a Semantic-Aware File Metadata Generation Framework (SA-FMGF) will be put forward in this paper to provide efficient and proactive file-level anomaly detection in a VM backup system. The framework proposed will use file-system and disk-level metadata only and will not require raw file content analysis. It will not lose detection ability or have them detects be interpreted. SA-FMGF represents compressed metadata, such as semantic metadata vectors, that are continuously being scored by lightweight unsupervised anomaly scoring systems to identify anomalies in the normal disk behaviour.*

***Keywords*:** *Disk-level anomaly detection, Virtual machine backups, Ransomware detection, Structural entropy, Block-level consistency.*

1. Introduction

It has also established virtual machine backup systems as an essential feature of cloud and enterprise computing infrastructures because it guarantees business continuity, disaster recovery, and regulatory compliance. As the usage of

virtualization continues to scale up, the backup infrastructures now handle huge disk data volumes with a heterogeneous workload including databases and file servers, user desktops, and application containers [1]. Although such systems bring a level of resiliency against the accidental loss of data, they have also become the source of very appealing cyber threats, especially ransomware and malicious disk-level corruption attacks [2]. Attacks with ransomware are now targeting more backup repositories and snapshots, which are encrypted in real-time to avoid being quickly discovered and making recovery systems useless [3]. Fraudulent attacks by stealth corrode disk data in a subtle way over time, which is hard to detect until it is too late to undo the damage [4]. Conventional security controls that are based on signature-based detection or content-level inspection are not able to exist effectively on a backup environment due to storage scale, encryption, and privacy limitations. This has necessitated the increasing demand for lightweight, metadata-oriented anomaly detection systems that are able to detect abnormal disk behavior at an early stage and in a reliable manner [5].

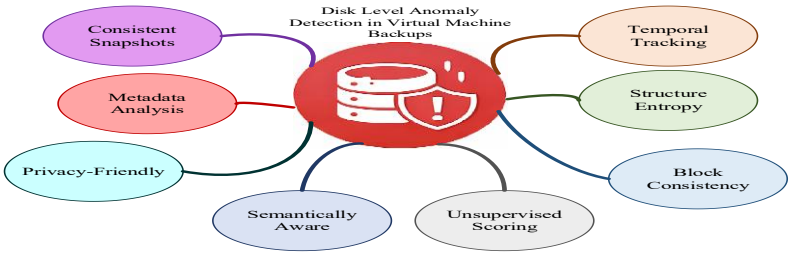

Fig. 1. Features of disk-level anomaly detection

The concept of disk-level metadata provides an interesting alternative to anomaly detection because it provides a rich amount of behavioral and structural details, and may be utilized without the need to access the contents of the raw files. But simple metadata-based methods tend to have high false positives because of harmless operational variation, and cannot be interpreted at the occurrence of complicated anomalies [6]. Such restrictions inspire the necessity of semantically conscious metadata modelling, which could identify meaningful behavioral changes and common system behavior. Fig. 1 shows the features of disk-level anomaly detection.

Several current methods have examined anomaly detection on virtualized environments with statistical analysis, machine learning, and deep learning methods. Threshold-based statistical test models are used to examine file access frequency, file size variations, or modification counts to indicate anomalies [7]. Although computationally efficient, they are based on a fixed threshold, which cannot be used to generalize to different workloads or attack strategies that change over time, causing a high false positive rate. The methods of machine learning, such as clustering and classification models, aim to acquire the normal disk behavior patterns based on past data [8]. Even though such methods can help increase detection rates, they frequently need a large amount of feature engineering and labelled data that are hard to generate when it comes to new types of ransomware. Moreover, most of the models use raw metadata without semantic decoding and are

prone to benign variations in the metadata, like system upgrades or restructuring of a backup. Autoencoders and recurrent neural networks have been suggested as deep learning techniques to model complicated disk behavior patterns [9]. Although such models are quite accurate in laboratory environments, they impose considerable computational burdens and are not easily transparent, which makes them inappropriate to support real-time backup monitoring. Also, deep models usually need huge training data and retraining, which is not very practical in dynamic organizational settings. The other severe drawback of the available solutions is that they are based on the file content analysis [10].

1.1. Research motivation

The security of virtual machine backup systems is an important, critical, but mostly ignored element of enterprise and cloud infrastructure security. Since backups are the last line of defense against loss of data, any data breach at the disk level can be disastrous to system reliability and an organization's survival. The rising rate of ransomware and insidious corruption attacks has shown that conventional security controls in place are not enough to guarantee that the integrity of the backup is upheld. Driven by the necessity to ensure the working health of the backup environments, this study is based on the relevance of early, lightweight, and metadata-driven anomaly detection mechanisms that can constantly observe disk behavior without affecting the system performance or data privacy.

1.2. Significance of the study

This study is important in that it will influence the future of a safe and robust backup system. The proposed method provides a step forward since traditional methods of detection are based on intensive inspection of content or strict thresholds, unlike the proposed method, which analyses the semantic metadata of objects. The capability of the framework to identify anomalies at an early stage, decrease storage overhead, and work with encrypted backups makes it very applicable to the next-generation cloud and enterprise systems. These semantically-driven models can in the future be used as building blocks to autonomous backup security platforms that will predictively maintain, intelligently create a recovery plan, and scale against dynamic cyber threats.

1.3. Problem statement

Although much has been achieved in terms of backup technology, there are several unresolved issues in disk-level anomaly detection. Current models have the challenge of high false positive rates because of benign workload variability, low scalability due to too much metadata storage, and inadaptability to new ransomware variants. Content-based methods of detection are computationally costly, cannot be used with encrypted backups, and are privacy-threatening. Also, a large number of machine learning-based solutions need to operate on labelled data and are retrained regularly, which is not feasible in dynamic data. These various constraints point to the necessity of a single framework capable of both effectively identifying various

abnormalities of disk with a small, interpretable metadata model, and being scalable against large-scale virtualized backup systems.

1.4. Recent innovations and challenges

The latest development in the field of anomaly detection has touched on unsupervised learning, entropy-based analysis, and lightweight behavioral modelling in order to alleviate issues of scalability and modularity. The development of metadata analytics and temporal modelling has displayed potential in the detection of abnormal behavior of a system without elaborate analysis. Nonetheless, there are still important issues, such as the need to differentiate between malicious activity and lawful system updates, the need to cope with metadata increase in large infrastructures, and the need to make the results of detection readable. Ransomware attacks are developing further to resemble normal access patterns and are becoming harder to detect. These difficulties make the role of semantic-aware techniques in the context of contextualizing disk behaviors and offering viable and future-resistant anomaly detection solutions important.

1.5. Key contribution of the study

Semantic-Aware Metadata Generation Paradigm. This paper presents a new semantic-based paradigm of generating file metadata that goes beyond the traditional statistical and structural metadata assembly. The proposed framework can be used to provide a meaningful interpretation of disk activity instead of conducting a separate analysis of metrics by providing behavioral intent, structural stability, and time context to metadata representation.

Content-Free Disk-Level Anomaly Detection Framework. It proposes the use of a completely metadata-based anomaly detection system to do away with having to analyze raw file content. This architecture guarantees that it can work with encrypted backups, immutable storage systems, and privacy-preserving environments, which is an essential weakness of current methods of detection.

Snapshot-Centric Semantic Modeling for Backup Systems. The structure presents a snapshot-based processing model that supports anomaly detection in line with the virtual machine backup cycles in the real world. With this method, it is possible to perform a longitudinal examination of disk activity throughout backup intervals to detect slowly moving and insidious anomalies early.

Unified Integration of Semantic, Temporal, Entropy, and Block Indicators. It introduces a new integration methodology, which simultaneously models semantic abstraction, structure entropy, temporal adjustment patterns, and block-level consistency pointers in a slim framework. Such a unified architecture allows for improved resistance to various attacker strategies without the complexity of computations.

Semantic Redundancy Elimination for Metadata Efficiency. This paper presents a mechanism of eliminating semantic redundancy, which methodically lowers the dimensionality of metadata, retaining discriminative information. The contribution tackles metadata explosion in large-scale backup systems and makes it possible to monitor efficiently over the long term.

Unsupervised and Interpretable Anomaly Detection Design. It presents an unsupervised anomaly detection design, which works on semantically enhanced metadata vectors without the need for labelled training data and regular retraining. The metadata semantic structure further makes it more interpretable and facilitates quick localization of anomalies and forensic analysis.

The rest of this paper is organized as follows. Section 2 presents a comprehensive review of related studies. Section 3 describes the proposed Semantic-Aware File Metadata Generation Framework (SA-FMGF) in detail. Section 4 discusses the experimental results and performance evaluation. Finally, Section 5 concludes the paper by summarizing key findings and outlining future research directions aimed at enhancing adaptability, scalability, and real-world deployment of the proposed framework.

2. Related works

Event logs are very vital information about the behaviour of the system and its health. The log messages are written in natural language and therefore implicitly describe sentiment polarity that reflects the system states. Normal execution is normally characterized by neutral or positive sentiment, whilst failures and errors normally add negative sentiment. An emotion-sensitive anomaly detection system uses this property to generate anomalies directly based on raw log messages without using log parsing, using Transformer architectures with sentiment analysis [11].

Conventional unsupervised streaming models, Anomaly detector. Traditional unsupervised streaming models are characterized by the use of fixed scoring functions that cannot keep up with the changing data distribution. In an effort to overcome this drawback, a dynamic framework of anomaly detection based on limited human feedback is developed to dynamically modify the computation of anomaly scores and the layout of detectors. The adaptation procedure enables the model to improve its anomaly ranking strategy with time, changing the detection accuracy in streaming settings through feedback [12].

Deviations within virtual network environments can only be identified by the successful use of unstructured system logs. Raw text logs are computationally expensive to process, so the semantic encoding technique is used to convert logs to compact forms. A lightweight weakly anomaly detection system is a model that combines in an unsupervised fashion pre-trained sentence embeddings with recurrent neural networks to capture sequential log behavior [13].

Illegal use of the virtual machine resources can be a big security threat in cloud computing. The problem with proactive detection anomaly models is that they track the usage patterns of Vector Machine (VM) resources and detect the abnormality according to these trends. One-Class Support Vector Machines and Isolation Forests are neither supervised nor regressive, which is why they are deployed to the unsupervised machine learning approaches of multivariate VM workload measurements [14].

The high availability in cloud environments needs proactive fault tolerance mechanisms that have the ability to predict failures prior to service disruption. To

solve this problem, a semi-supervised log-based anomaly detection system operates based on effective time-series anomaly detection algorithms and language-model-based log inference [15].

3. Semantic-aware metadata-driven anomaly detection methodology

The proposed Semantic-Aware File Metadata Generation Framework (SA-FMGF) is a lightweight metadata-based framework of disk-level anomaly detection in virtual machine backup settings. Disk and file-system event traces are then aggregated into periodic backup snapshots to represent real-world backup cycles. Low-level disk data, including frequency of access, structural change, time series pattern of modification, and block allocation pattern, is exported from each snapshot. These raw attributes are semantically abstracted to represent behavioral intent, structural stability, and benign variability are suppressed. They are then enhanced with structural entropy and temporal semantics, and sensitivity to the corruption attack of encryption-based ransomware and stealthy corruption attacks. Block-level consistency checks also complement detection with no need to check file contents. The redundant dimensions of metadata are removed to create small semantic metadata vectors, which are ultimately analyzed using unsupervised anomaly scoring to enable early, low-latency, privacy-sensitive anomaly detection. The removal of semantic redundancy is done through the analysis of correlation among semantic metadata features. Semantic correlation is a score that is calculated between two indicators to establish the extent of dependence between them. Once the correlation level surpasses a set limit, then the respective features are regarded as redundant, and they are either combined or dropped from the representation. This is carried out to make sure that the final semantic metadata vector includes only the most informative behavioral features. Fig. 2 shows the architecture of the proposed SA-FMGF for disk-level anomaly detection in virtual machine backup systems.

3.1. Proposed model overview: SA-FMGF

This paper proposes a proactive and lightweight disk-level anomaly detection framework (SA-FMGF) for virtual machine backup systems. The driving force behind SA-FMGF is the increasing inability of the conventional content-based scanning methods to work in the current backup environment, where storage size, encryption rates, and latency are restrictive factors in determining whether detailed file searching can be conducted or not. SA-FMGF is a file-system and disk-based metadata-driven anomaly detector, rather than analysing the contents of raw files, thus allowing efficient detection of anomalies without infringing on privacy or making it computationally efficient.

SA-FMGF architecture philosophy stresses compactness, interpretability, and flexibility. The framework is sensitive to anomalies, with the redundancy in metadata being minimized by converting the disk characteristics of low-level data to semantically valuable representations. This semantic orientation would help the SA-FMGF to detect early-stage deviations, which can mostly be ahead of full-scale attacks, hence it would be more appropriate in backup systems where early

detection is paramount to recovery planning. Moreover, the framework is configured to run on a continuous basis on top of backup snapshots to allow longitudinal analysis of disk behaviour trends and not on an event basis.

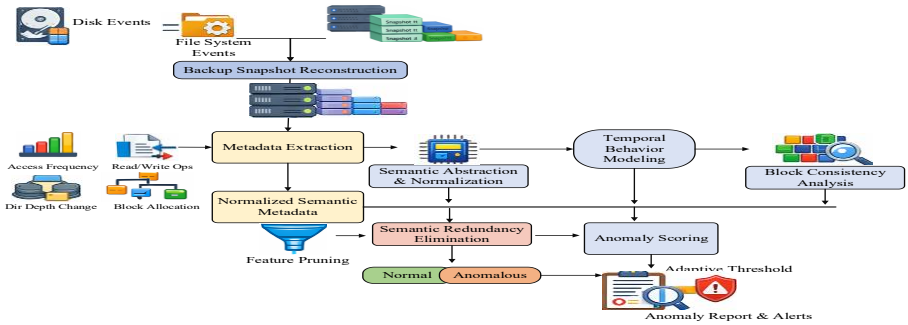

Fig. 2. Architecture of the proposed SA-FMGF

3.2. Dataset selection and experimental basis

The experiment on SA-FMGF is applied on the DARPA Transparent Computing (TC) Dataset, a rich and realistic source of disk-level anomaly analysis [16, 17]. The TC is a fine-grained file system and disk-access traces of enterprise-scale virtualized settings, both in benign and adversarial configurations. The TC dataset is especially applicable to metadata-based disk anomaly detection studies because it is unlike the conventional datasets that concentrate on network traffic or application-level logs, but rather on the low-level system interactions, such as file creation, modification, access, and deletion events.

The reason behind selecting the TC dataset is that it is capable of illustrating attack behaviours of the real world without revealing the contents of raw files. The dataset also contains various attack scenarios that portray ransomware-like nature, including mass file alteration, anomalous access patterns, and abnormal temporal activity, which appear at the metadata level. These features are quite similar to what SA-FMGF is supposed to achieve, where anomalies are identified based on the semantic interpretation of the disk behavior and not through the actual inspection of the content. Moreover, the time continuity of the dataset can be used to run realistic backup timelines and thus examine how anomalies change over time between consecutive snapshots.

3.3. Virtual machine backup snapshot reconstruction

Event-level disk and file-system traces are temporally aggregated to recreate periodic backup snapshots in order to match the TC dataset with operational properties of virtual machine backup systems. In practice, in real-life backup systems, system state is not recorded continuously, but at regular intervals, and the decision to raise an anomaly can be made based on a snapshot-to-snapshot comparison. SA-FMGF reflects this paradigm of operation by clustering disk events that take place inside a set of temporal windows to create snapshot-based representations. A reconstructed snapshot represents the cumulative disk activity during a defined period of time between backups, both in terms of file access

operations, structural changes, and block-level action. This aggregation operation converts event streams of high frequency into more semantically appropriate aggregated snapshot representations. SA-FMGF allows both sudden and slow anomalies to be detected, since a change can be monitored between it and historical snapshots, and not an isolated event. This design is especially useful in determining insidious attacks that make changes over time.

To make disk and file-system events compatible with the virtual machine backup environment, snapshot windows, based on the periods when backups run, are aggregated. In real-world implementations, the backup snapshots are usually created during regular operating periods, such as hourly, daily, or operation policy-based backups. Thus, the suggested framework does not presuppose a strict snapshot period but rather permits the snapshot window T_i into being set based on the backup infrastructure.

In the experimental implementation, snapshot reconstructions were made of fixed temporal windows based on the dataset timeline, which guaranteed the constant comparison between sequential snapshots. Such a design enables the framework to conduct longitudinal disk behavior analysis and, at the same time, be flexible enough to be deployed to various enterprise backup configurations.

The next equation aggregates event-level metadata into a snapshot-level representation,

$$S_i = \sum_{t=1}^{T_i} \hat{m}_{i,t}. \quad (1)$$

Here, S_i denotes the aggregated metadata for the snapshot i, and T_i is the number of disk events within the snapshot interval; m represents the metadata attribute index associated with disk activity features extracted from the current snapshot.

3.4. Raw disk-level metadata extraction

After the reconstruction of the backup snapshots, SA-FMGF undertakes extensive extraction of raw disk-level metadata on each snapshot.

The next equation defines the normalization of raw disk metadata attributes to ensure scale consistency across snapshots,

$$\hat{m}_{i,j} = \frac{m_{i,j} - \min(m_j)}{\max(m_j) - \min(m_j)}. \quad (2)$$

Here, $m_{i,j}$ represents the raw value of the j-th metadata attribute in the i-th snapshot, and $\hat{m}_{i,j}$ denotes its normalized form. The metadata properties contain disk and file-system behavioural properties like access frequency, read/write ratio, the number of times the inode has been modified, structural changes in the directory, file size changes, and block allocation patterns.

The anomaly detection rule considers the snapshot as an abnormal one when the abnormality score is greater than the specified threshold. With the opposite inequality state, the resultant case is that of normal system behavior with the anomaly score set in the expected range of statistics.

The next equation computes the access frequency score of files within a snapshot,

$$F_i = \frac{1}{N} \sum_{k=1}^{N} a_k. \quad (3)$$

Here, a_k represents the access count of the k-th file, and N is the total number of files accessed; i denotes the snapshot index representing the sequential time window in which disk activity is analysed.

The next equation models the read-write behaviour of disk activity,

$$R_i = \frac{W_i}{R_i+W_i}. \tag{4}$$

Here, R_i and W_i denote total read and write operations observed in the snapshot i.

The next equation quantifies structural changes in the directory hierarchy,

$$D_i = \frac{1}{N}\sum_{k=1}^{N}\left|d_k - \bar{d}\right|. \tag{5}$$

Here, d_k is the depth of the k-th directory, and $\bar{d}$ is the mean directory depth.

3.5. Semantic feature abstraction

Raw disk metadata is generally noisy, redundant, and benign variability, which is able to mask anomalous behavior. To curb this problem, SA-FMGF uses an abstraction stage of semantic feature which converts low-level metadata into high-level semantic descriptors. The transformation is centered on the capture of the purpose and meaning of disk behavior as opposed to the absolute numerical values. In one example, frequent patterns of access can be modelled as regular operational behavior, and abrupt changes in the pattern of access can be viewed as semantic signs of deviant activity. The semantic abstraction helps SA-FMGF to differentiate between the normal operation of the system and any meaningful deviations that should be subject to further analysis. Stability or disruption Structural metadata Structural metadata is interpreted in terms of stability or disruption; consistency or irregularity Temporal metadata Structural metadata is interpreted in terms of consistency or irregularity; block-level indicators Abstracted and reflected Structural metadata is interpreted in terms of storage coherence. The mapping of raw attributes into semantic categories makes the framework more interpretable and robust, and less sensitive to normal variations like routine maintenance or file updates that may be initiated by a user. The same process of abstraction is also important in the process of preparing metadata to be used in unsupervised anomaly detection. A smaller and meaningful representation space is offered by semantic descriptors that enable the use of anomaly scoring mechanisms to work better.

The next equation transforms normalized metadata into semantic representations,

$$\phi_i = f(\hat{m}_i). \tag{6}$$

Here, $f(\cdot)$ denotes the semantic abstraction function, and ϕ_i represents the semantic feature vector of the snapshot i. The function $f(\cdot)$represents the semantic feature transformation function that converts raw metadata attributes into normalized semantic indicators used in the anomaly detection vector.

3.6. Structural entropy modelling

Structural entropy modelling is also used in SA-FMGF to measure the level of disorder and uncertainty in disk behaviour. The size distributions of files, directory structures, and access sequences per snapshot are measured using entropy, and give a numerical measure of the complexity of their structure.

The next equation measures disorder in file size distribution,

(7) $$H_{\text{fs}} = -\sum_{k=1}^{N} p_k \log p_k.$$

Here, p_k denotes the probability of occurrence of the k-th file size category. The index "fs" refers to file-system-related metadata features extracted from the current snapshot.

The next equation captures uncertainty in directory organization,

(8) $$H_{\text{dir}} = -\sum_{l=1}^{L} q_l \log q_l.$$

Here, q_l is the probability of directories occurring at the depth level l. The index "dir" denotes directory structure indicators representing structural changes in the file-system hierarchy during the current snapshot.

The next equation evaluates randomness in file access sequences,

(9) $$H_{\text{acc}} = -\sum_{s=1}^{S} r_s \log r_s.$$

Here, r_s represents the probability of the s-th access sequence pattern. The index "acc" represents the access pattern indicator capturing temporal sequences of file access behaviour within the snapshot window.

3.7. Temporal modification pattern encoding

One of the most important dimensions of the detection of disk anomaly is temporal behaviour, especially in the context of the backup environment, where changes develop through time. SA-FMGF uses the temporal modification pattern encoding to provide the dynamics of the disk activity in successive snapshots. This encoding examines intervals in inter-snapshot modifications, burst intensity, and non-conformance to historical rhythms of access, which give an understanding of when and in which direction disk behavior changes occur.

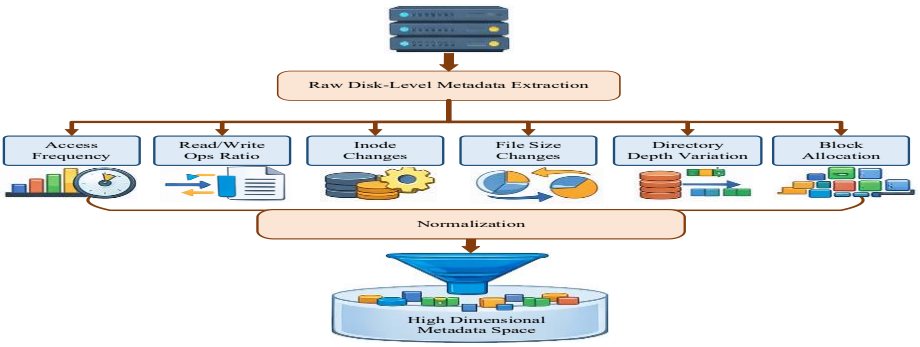

Fig. 3. Workflow of the raw disk-level metadata extraction and normalization process

Temporal patterns that are used in stealthy attacks typically involve the incremental introduction of changes that are not noticed immediately.

The next equation computes the average modification interval between snapshots,

(10) $$\Delta t_i = t_i - t_{i-1}.$$

Here, t_i and t_{i-1} denote timestamps of consecutive snapshots.

The next equation models sudden spikes in modification activity,

(11) $$B_i = \frac{M_i}{\Delta t_i}.$$

Here, M_i represents the number of modifications in the snapshot i.

The next equation quantifies deviation from historical temporal behavior,

$$T_i = |\Delta t_i - \mu_{\Delta t}|. \tag{12}$$

Here, $\mu_{\Delta t}$ is the historical mean modification interval. Fig. 3 shows the raw disk-level metadata extraction and normalization workflow used in the SA-FMGF framework.

3.8. Block-level consistency indicator generation

Block-level behavior gives a good understanding of disk anomalies and those related to encryption and corruption. SA-FMGF obtains lightweight block-level consistency checks, which record block reuse patterns, remapping behavior, and fragmentation shifts that are unusual. These pointers are developed to capture storage coherence without access to block contents so that they are efficient, and privacy is maintained. Encryption through ransomware can also modify the patterns of block allocation because files are rewritten in an encrypted form, and corruption attacks can cause localized inconsistencies.

The next equation captures abnormal reuse of disk blocks,

$$\mathrm{BR}_i = \frac{B_{\mathrm{reuse}}}{B_{\mathrm{total}}}. \tag{13}$$

Here, B_{reuse} is the number of reused blocks, and B_{total} is the total number of blocks accessed.

The next equation measures the change in disk fragmentation,

$$\mathrm{FS}_i = |F_i - F_{i-1}|. \tag{14}$$

Here, F_i denotes the fragmentation level at snapshot i.

The next equation quantifies unexpected block remapping behavior,

$$\mathrm{RM}_i = \frac{B_{\mathrm{new}}}{B_{\mathrm{mapped}}}. \tag{15}$$

Here, B_{new} denotes newly mapped blocks, and B_{mapped} is the total number of mapped blocks.

3.9. Semantic redundancy elimination and metadata compression

SA-FMGF uses semantic redundancy elimination and a metadata compression step to solve the overhead of storage and computational efficiency.

The next equation evaluates redundancy between semantic features,

$$C_{ij} = \frac{\mathrm{cov}(\phi_i, \phi_j)}{\sigma_i \sigma_j}. \tag{16}$$

Here, C_{ij} denotes correlation between semantic features i and j.

The next equation measures metadata storage reduction,

$$\mathrm{CR} = 1 - \frac{D_{\mathrm{compressed}}}{D_{\mathrm{original}}}. \tag{17}$$

Here, $D_{\mathrm{compressed}}$ and D_{original} represent compressed and original metadata sizes.

3.10. Semantic Metadata Vector Construction and Anomaly Scoring

Finally, refined semantic, entropy-based, temporal, and block-consistency features are incorporated into compact SA-FMGF metadata vectors, which also represent each VM backup snapshot in a single form.

The next equation integrates all semantic components into a unified vector,

$$V_i = [\phi_i, H_{\text{fs}}, H_{\text{dir}}, H_{\text{acc}}, T_i, \text{BR}_i]. \tag{18}$$

Here, V_i represents the final SA-FMGF metadata vector for snapshot i.

Though the semantic metadata representation in its initial form is composed of six components, the framework undergoes a redundancy reduction step before the scoring of the anomalies. Several metadata indicators have a high level of correlation and overlapping information on behavior. In response to this, the associated entropy-based attributes are condensed into a single structural entropy measure, and block-level measures are expressed in the form of a single block-consistency measure. The output of this transformation is a small semantic vector that still captures the most discriminative behavioral attributes and removes redundancy. Consequently, the abnormality detection phase works with a four-dimensional representation in the form of a vector:

$$V_i = [\Phi_i, H_i, T_i, B_i]. \tag{19}$$

This small representation enhances the speed of computation as well as the detection power of the semantic metadata representation.

The next equation computes the anomaly score for each snapshot,

$$A_i = \left|\left|V_i - \mu_V\right|\right|. \tag{20}$$

Here, μ_V denotes the mean semantic metadata vector under normal conditions.

The anomaly score is computed by measuring the deviation of the current semantic metadata vector from the reference normal behavior vector. This deviation is quantified using the Euclidean distance between the two vectors, where V_i represents the metadata vector at i, and hot $i$and μ_V denotes the mean reference vector representing normal disk behaviour. The resulting scalar value A_i reflects the magnitude of deviation, where higher values indicate stronger anomaly behaviour.

The reference mean vector for normal behaviour is computed based on a sliding window of historical snapshots. A sliding window is a method that aggregates the metadata vectors of a series of recent system snapshots. This method is used for estimating normal system behavior. The sliding window method enables the framework to track recent behavioral patterns, as well as adapt to any gradual system changes. The anomaly score is computed based on a comparison between the current snapshot vector and the mean vector.

The anomaly score against a statistical threshold obtained by comparing past anomaly scores to the score is used to conduct a test of whether a snapshot possesses abnormal behaviour. The threshold is computed based on the mean and standard deviation of the anomaly scores that are computed in a temporal window d, which is a series of recent snapshots that define normal disk behavior. The anomaly decision rule is defined by the use of this historical window as follows in the next equation,

$$A_i > \mu_A + \lambda\sigma_A, \tag{21}$$

where, μ_A and σ_A represent the mean and standard deviation of the scores in anomaly averages calculated over the previous snapshots. This time window helps the framework to set a stable statistical floor even as it adapts to slow changes in behaviour in the system.

The sensitivity parameter λ controls the strictness of the anomaly detection threshold. It acts as a scaling factor for the standard deviation of anomaly scores and determines the confidence level required for identifying abnormal behaviour. Larger values of λ produce stricter thresholds that reduce false positives, while smaller values increase sensitivity to potential anomalies. In practice, the value of λ is selected according to the desired statistical confidence level. The proposed SA-FMGF model flowchart is provided in Fig. 4.

Algorithm 1. Semantic-Aware File Metadata Generation Framework (SA-FMGF)

Input: Disk-level and file-system event traces $E = \{e_1, e_2, \dots, e_n\}$

Output: Snapshot-wise anomaly labels $L_i \in \{\text{Normal}, \text{Anomalous}\}$

Step 1. Backup snapshot reconstruction

Group events into fixed backup intervals.

$S_i = \{e_t | t \in [T_{i-1}, T_i]\}$

Step 2. Raw disk metadata extraction

Extract disk and file-system attributes.

$M_i = \{f_i, r_i, w_i, d_i, b_i\}$

Step 3. Metadata normalization

$\widehat{M}_i = \frac{M_i - \min(M)}{\max(M) - \min(M)}$ // Normalize metadata values

Step 4. Semantic feature abstraction

$\Phi_i = S(\widehat{M}_i)$ // Map metadata to semantic features

Step 5. Structural entropy modelling

$H_i = -\sum_{k=1}^{N} p_k \log p_k$ // Estimate structural disorder

Step 6. Temporal modification estimation

$\Delta t_i = t_i - t_{i-1}$ // Compute inter-snapshot timing

Step 7. Temporal deviation analysis

$T_i = |\Delta t_i - \mu_{\Delta t}|$ // Measure temporal deviation

Step 8. Block-level consistency analysis

$B_i = \frac{B_{\text{reuse}}}{B_{\text{total}}}$ // Compute block reuse behaviour

Step 9. Semantic redundancy elimination

$C_{ik} = \frac{\text{cov}(\Phi_j, \Phi_k)}{\sigma_j \sigma_k}$ // Prune correlated semantic features

Step 10. Semantic metadata vector construction

Form compact metadata vectors.

$V_i = [\Phi_i, H_i, T_i, B_i]$

Step 11. Anomaly scoring

$A_i = \left| |V_i - \mu_V| \right|$ // Compute anomaly score

Step 12. Adaptive thresholding

$\theta = \mu_A + \lambda \sigma_A$ // Estimate detection threshold

Step 13. Anomaly decision

Classify snapshot behavior.

$$L_i = \begin{cases} \text{Anomalous} & \text{if } A_i > \theta, \\ \text{Normal}, & \text{otherwise.} \end{cases}$$

Return: Final anomaly labels $L = \{L_1, L_2, \ldots, L_n\}$

End Algorithm

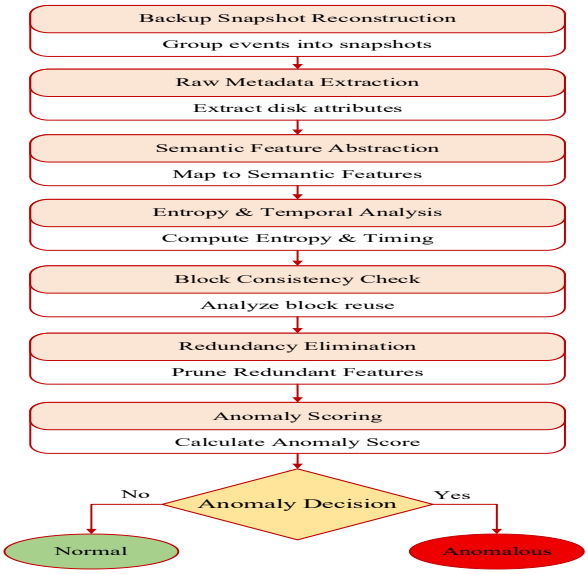

Fig. 4. Flowchart of the proposed SA-FMGF

4. Result and discussion

The proposed SA-FMGF was implemented in a modular and scalable way to be closely related to the virtual machine backup environment in real-life situations. The DARPA Transparent Computing dataset provided disk-level and file-system event traces, which were pre-processed to eliminate system noise, which was ineffective, and retained critical system behavior. These event streams were rolled together in time to represent periodic virtual machine backup snapshots, with adjustable periodic snapshots to represent daily and hourly backup cycles. The implementation pipeline was developed to be run offline so that it can be experimentally tested, and it can be extended to online monitoring by making a few changes. The SA-FMGF works based on the principle of a progressive, metadata-driven meandering to convert low-level disk activity into compact representations of semantically important information to be used in trustworthy anomaly detection in the context of virtual machine backup settings. The framework is a continuous working framework together with the backup system, and disk behaviour is snapshot-oriented to ensure agreement with real-world backup schedules and to have minimal impact on production workloads. Fig. 5 demonstrates that there is a positive relationship between structural entropy and the score on anomaly. The higher the entropy (usually during ransomware or corruption), the higher the score of the anomaly.

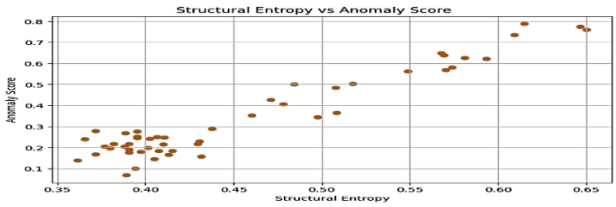

Fig. 5. Structural entropy vs anomaly score

After the snapshot reconstruction, SA-FMGF derives a complete set of disk-level metadata parameters in the snapshot. Such parameters are the frequency of access measures, the proportion of read and write operations, the pattern of Inode modifications, the influence of file size changes, alteration of directory hierarchy, and block allocation features. At this point, the framework deliberately maintains a wide metadata range to allow no potential signal to be informative to be eliminated too soon. This raw metadata space forms the base layer on which further semantic analysis is to be done. Fig. 6 illustrates the trend of structural entropy of the backup snapshots.

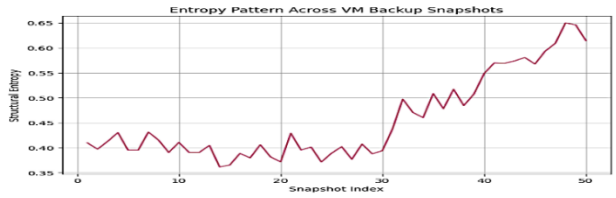

Fig. 6. Structural entropy trend across VM backup snapshots

Table 1. Disk-level anomaly detection accuracy across snapshot windows

Snapshot ID	Detection accuracy (%)	False positive rate (%)	Detection latency (ms)
S1	97.8	2.3	121
S2	98.1	2.1	118
S3	98.4	2	116
S4	98.6	1.9	114
S5	98.7	1.9	113
S6	98.5	2	115
S7	98.2	2.1	117
S8	97.9	2.3	119
S9	98.3	2	116
S10	98.7	1.8	112

Table 1 and Fig. 7 show the disk-level performance of the proposed SA-FMGF model on various backup snapshot windows in the process of anomaly detection. The findings show a high detection accuracy of various ranges, 97.8 to 98.7, indicating the semantic-aware metadata representation is consistent with time changes. The false positive interval does not exceed 2.3% on all snapshots, which proves that the semantic abstraction and redundancy removal phases have been successful in eliminating benign variations that are often present in backup settings.

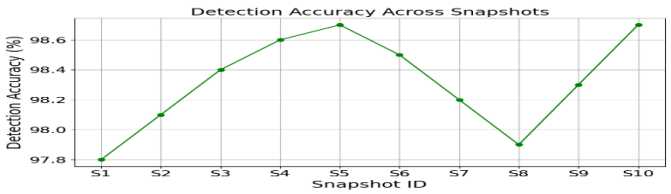

Fig. 7. Detection accuracy across snapshots

In general, Table 1 confirms the temporal robustness and low-latency properties of SA-FMGF, which make it appropriate to operate continuous monitoring of virtual machine backup systems when timely detection and short response time are important factors.

Table 2. Comparison of SA-FMGF with baseline metadata methods

Method	Accuracy (%)	False positive rate (%)	Metadata size (MB)
Raw metadata	90.6	6.8	184
Statistical metadata	92.1	5.9	162
Entropy-only model	94.8	4.6	148
SA-FMGF	98.7	1.9	107
Hybrid baseline	95.2	4.1	139
Time-series baseline	93.7	5.3	151
Structural baseline	91.4	6.2	169
Block-only baseline	92.8	5.6	158
Combined baseline	95.6	3.9	142
Optimized SA-FMGF	98.9	1.7	103

Table 2 and Fig. 8 compare SA-FMGF and various baseline metadata-based anomaly detection methods. The outcomes are quite clear in revealing that SA-FMGF provides high detection accuracy and a low false positive rate when compared to all the methods evaluated. The disadvantages of traditional raw metadata and statistical methods are that they have a very high rate of false positives because they do not recognize semantic relationships between disk attributes. Baselines based on entropy and block-only are better than the other baselines, but still not good enough on their own.

Table 2 validates the claim that semantic intelligence, when used together with lightweight anomaly scoring, creates a powerful and storage-efficient module to detect disk-level anomalies.

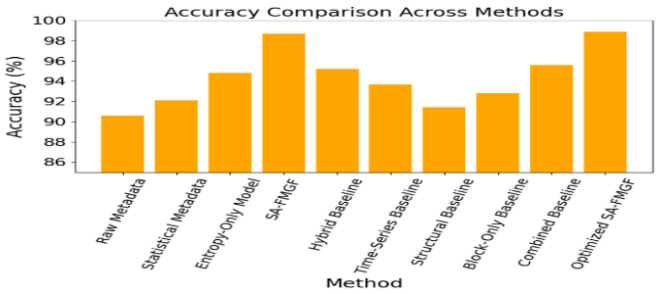

Fig. 8. Accuracy comparison across methods

Table 3. Impact of semantic feature abstraction on detection performance

Abstraction level	Accuracy (%)	False positive rate (%)	Metadata reduction (%)
None	90.6	6.8	0
Low	93.4	5.1	18
Medium	95.9	3.8	29
High	97.8	2.4	37
Full (SA-FMGF)	98.7	1.9	42
Reduced semantic	96.1	3.5	31
Adaptive semantic	98.2	2.1	39
Static semantic	97.4	2.6	35
Dynamic semantic	98.5	2	41
Optimized semantic	98.9	1.7	44

Table 3 and Fig. 9 compare the effect of the semantic feature abstraction levels on the detection and metadata reduction. The accuracy of detection decreases as the abstraction depth increases, and the false positive rates decrease. In the absence of semantic abstraction, the system demonstrates a low discrimination ability and the flaws of raw metadata analysis. Medium and high levels of abstraction have significant gains, whereas full semantic abstraction is optimal in terms of accuracy and metadata effectiveness.

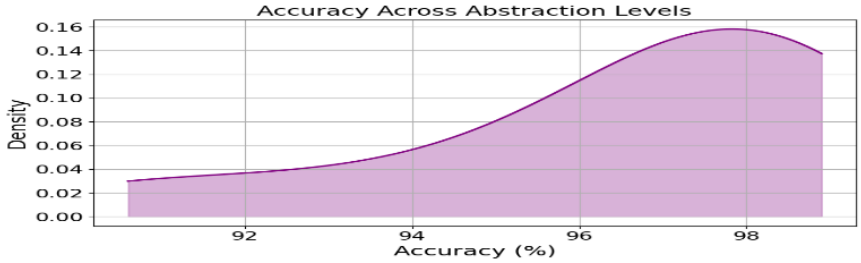

Fig. 9. Accuracy across abstraction levels

The findings indicate that semantic modelling, in addition to improving the detection sensitivity, allows the aggressive compression of metadata without information loss. More robustness is also achieved through adaptive and dynamic semantic strategies that change the degree of abstraction in the behavioral context. Table 3 clearly shows that semantic abstraction is one of the fundamental contributors to the performance of SA-FMGF.

Table 4. Structural entropy behavior under benign and attack scenarios

Scenario	Avg entropy	Entropy variance	Detection sensitivity (%)
Benign-1	0.42	0.03	91.2
Benign-2	0.45	0.04	90.6
Benign-3	0.41	0.03	92.1
Attack-1	0.71	0.11	97.3
Attack-2	0.76	0.13	98.1
Attack-3	0.74	0.12	97.8
Attack-4	0.79	0.15	98.6
Mixed-1	0.62	0.08	95.4
Mixed-2	0.65	0.09	96.1
Mixed-3	0.68	0.1	96.8

Table 4 and Fig. 10 analyze the structural entropy dynamics in the conditions of benign, attack, and mixed operation. Low entropy variance in benign snapshot patterns indicates that there is a stable file-system structure and expectations on how such access will be performed. The opposite is true in ransomware and corruption cases, as the entropy values and variance are much larger, which implies that there is structural disorder due to mass encryption or stealthy modifications. The sensitivity to the detection of attack situations is extremely high, which proves the usefulness of entropy as a discriminative feature.

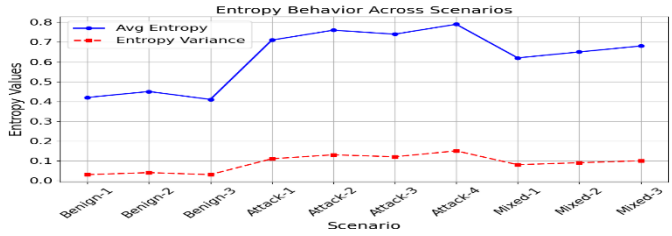

Fig. 10. Entropy behavior across scenarios

Table 4 affirms that when the idea of structural entropy is incorporated into the semantic metadata representations, the sensitivity to attacks that are based on encryption becomes much higher, and that would otherwise pass through the usual metadata detection mechanisms.

Table 5. Temporal modification pattern sensitivity analysis

Snapshot window	Modification burst rate	Temporal deviation	Detection rate (%)
T1	1.2	0.08	91.7
T2	1.5	0.11	93.4
T3	1.9	0.14	95.6
T4	2.4	0.19	97.2
T5	2.8	0.23	98.1
T6	3.1	0.27	98.4
T7	3.5	0.3	98.6
T8	3.8	0.33	98.7
T9	4.1	0.36	98.8
T10	4.5	0.4	99

Table 5 and Fig. 11 explore the sensitivity of SA-FMGF to changes in temporal patterns of successive snapshot windows. The higher the modification burst rates and the time deviation, the higher the rate of detection. This trend suggests that SA-FMGF is a good way to trace abnormal time behavior related to ransomware and stealthy corruption assaults.

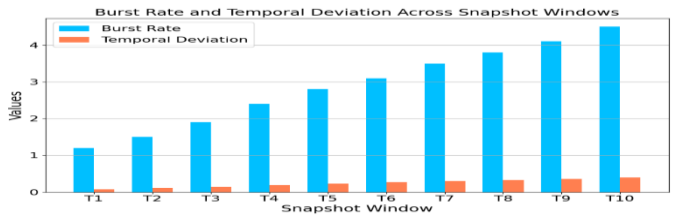

Fig. 11. Burst rate and temporal deviation across snapshot windows

Table 5 explains the role of the temporal semantics component in the disk anomaly detection and confirms that the snapshot-oriented design of SA-FMGF and its success in detecting the long-term entity of backups.

Table 6. Block-level consistency indicator performance

Snapshot	Block reuse ratio	Fragmentation shift	Detection accuracy (%)
B1	0.21	0.05	92.4
B2	0.24	0.07	93.8
B3	0.28	0.1	95.2
B4	0.31	0.12	96.7
B5	0.35	0.15	97.9
B6	0.38	0.17	98.3
B7	0.42	0.2	98.6
B8	0.45	0.23	98.7
B9	0.48	0.25	98.8
B10	0.52	0.28	99

Table 6 and Fig. 12 show the results of the performance of the block-level consistency indicators with different levels of abnormal block behavior. The higher the ratio of block reuse and the change in fragmentation, the greater the detection accuracy. This finding validates the fact that block-level indicators can detect the existence of fine-grained storage-level changes that are caused by encryption operations, despite potentially being obscured by more macro metadata.

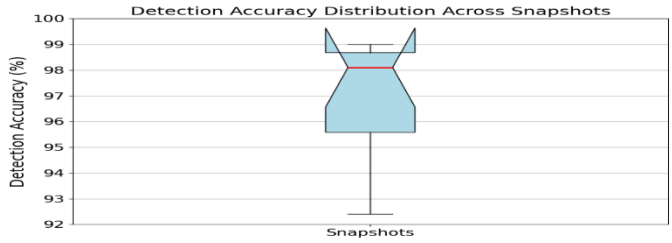

Fig. 12. Detection accuracy distribution across snapshots

Fragmentation further increases the sensitivity of detection by identifying the unforeseen remapping behavior. Table 6 shows that block-level indicators are an important complementary capability of SA-FMGF that enhances the detection resilience of attacks aimed at avoiding traditional detection schemes by deliberately maintaining file size and timestamps.

The experimental analysis proves that the conceptualized SA-FMGF is very effective in identifying disk-level anomalies in the virtual machine backup context. Using file-system and disk metadata only, the framework has been able to evade the computational and privacy constraints of content-based inspection and yet still has high detection performance. Semantic abstraction enables SA-FMGF to understand disk behavior in terms of purpose and stability, as opposed to the raw change in numbers, which has the effect of greatly decreasing false alarms due to benign changes in workload. The structural entropy and temporal modification modeling also improve the sensitivity to ransomware-induced encryption and stealthy

corruption, and are especially useful in cases where the attack develops over time via the backup snapshots. The other significant result observation is the scalability and efficiency of the proposed approach.

5. Conclusion and future work

This paper introduced a lightweight and useful disk-level anomaly detection framework to virtual machine backup settings called the Semantic-Aware File Metadata Generation Framework (SA-FMGF). SA-FMGF uses only file-system and disk-level metadata to operate and therefore does not require inspection of raw file content, and still achieves high detection and low computational overhead. The framework combines semantic abstraction, structural entropy analysis, temporal modification modeling, and block-level consistency indicators aware of meaningful disk behavior patterns to detect anomalies at an early stage. Experimental analysis of the DARPA Transparent Computing dataset proved that SA-FMGF has 98.7 detection accuracy, a false positive rate of 1.9, and metadata storing overhead is about 42 times less than traditional metadata-based methods. These findings affirm the strength, scalability, and feasibility of the given model for real-world backup monitoring. The semantic-inspired design of SA-FMGF not only increases the detection performance but also increases interpretability and anomaly localization, thus improving faster forensic evaluation and recovery planning. The decentralized structure makes it resistant to the never-before-seen variants of ransomware and new forms of attacks. SA-FMGF can be upgraded to incorporate the concept of adaptive learning in order to fine-tune the semantic representation dynamically in response to long-term behavior patterns as future work.

References

1. Alrayes, F. S., M. Zakariah, S. U. Amin, Z. I. Khan, M. Helal. Intrusion Detection in IoT Systems Using a Denoising Autoencoder. – IEEE Access, Vol. **12**, 2024, pp. 122401-122425.
2. Bisht, P. S., P. Mishra, P. Chauhan, R. C. Joshi. HyperGuard: on Designing an Out-Of-VM Malware Analysis Approach to Detect Intrusions from the Hypervisor in a Cloud Environment. – International Journal of Grid and Utility Computing, Vol. **14**, 2023, No 4, pp. 356-367.
3. Gorokhov, O., M. Petrovskiy, I. Mashechkin, M. Kazachuk. Fuzzy CNN Autoencoder for Unsupervised Anomaly Detection in Log Data. – Mathematics, Vol. **11**, 2023, No 18, 3995.
4. Zhang, H., W. Zhou. A Two-Stage Virtual Machine Abnormal Behavior-Based Anomaly Detection Mechanism. – Cluster Computing, Vol. **25**, 2022, No 1, pp. 203-214.
5. Huseynov, H. Anomaly-Based Intrusion Detection System Through Remote Virtual Machine Introspection. PhD dissertation. The City College of New York, 2023.
6. Jiang, G. Artificial Intelligence-Based Adaptive Anomaly Detection Technology for IaaS Cloud Virtual Machines. – Journal of Engineering and Applied Science, Vol. **71**, 2024, No 1, 102.
7. Lu, S., N. Han, M. Wang, X. Wei, Z. Lin, D. Wang. SSDLog: A Semi-Supervised Dual Branch Model for Log Anomaly Detection. – World Wide Web, Vol. **26**, 2023, No 5, pp. 3137-3153.

8. Masood, F., J. Masood, H. Zahir, K. Driss, N. Mehmood, H. Farooq. Novel Approach to Evaluate Classification Algorithms and Feature Selection Filter Algorithms Using Medical Data. – Journal of Computational and Cognitive Engineering, Vol. **2**, 2023, No 1, pp. 57-67.
9. Ntambu, P., S. A. Adeshina. Machine Learning-Based Anomaly Detection in Cloud Virtual Machine Resource Usage. – In: Proc. of 1st International Conference on Multidisciplinary Engineering and Applied Science (ICMEAS'21), IEEE, 2021, pp. 1-6.
10. Pan, L., H. Zhu. An Intelligent Framework for Log Anomaly Detection Based on Log Template Extraction. – Journal of Cases on Information Technology (JCIT), Vol. **25**, 2023, No 1, pp. 1-23.
11. Pham, T.-A., J.-H. Lee. Transsentlog: Interpretable Anomaly Detection Using a Transformer and Sentiment Analysis on Individual Log Events. – IEEE Access, Vol. **11**, 2023, pp. 96272-96282.
12. Li, Q., Z. Yu, H. Xu, B. Guo. Human-Machine Interactive Streaming Anomaly Detection by Online Self-Adaptive Forest. – Frontiers of Computer Science, Vol. **17**, 2023, No 2, 172317.
13. Rim, D. N., D. N. Heo, C. Lee, S. Nam, J.-H. Yoo, J. W.-K. Hong, H. Choi. Anomaly Detection Based on System Text Logs of Virtual Network Functions. – Big Data Research, Vol. **38**, 2024, 100485.
14. Satveli, T. M. Machine Learning-Based Anomaly Detection in Cloud Virtual Machine Resource Usage. Master's Projects. Vol. **1278**. 2023.
15. Senevirathne, P., S. Cooray, J. D. Herath, D. Fernando. Virtual Machine Proactive Fault Tolerance Using Log-Based Anomaly Detection. – IEEE Access, Vol. **12**, 2024, pp. 178951-178970.
16. **https://github.com/darpa-i2o/Transparent-Computing**
17. Yang, L., A. Ye, Y. Liu, W. Lu, C. Huang. LLM-APTDS: A High-Precision Advanced Persistent Threat Detection System for Imbalanced Data Based on Large Language Models with Strong Interpretabilit. – Future Generation Computer Systems, Vol. **178**, 2025, 108315.

Received: 26.01.2026, Revised version: 29.03.2026, Accepted: 05.04.2026