iJRASET
International Journal For Research in Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 14 **Issue:** V **Month of publication:** May 2026

**DOI:** https://doi.org/10.22214/ijraset.2026.83246

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 14 Issue V May 2026- Available at www.ijraset.com*

# AI Driven Insider Threat Detection Using Behavioural Analysis

Aishwariya Vijay Raj Laxmi, Swastika Srivastava
*Galgotias University*

***Abstract*—Insider threats represent one of the most critical challenges in modern cybersecurity. These threats arise from individualswithinanorganizationwhomisusetheirlegitimateaccess to harm the organization's assets, data, or operations. Traditional security mechanisms, primarily designed for external attackers, fall short in identifying these subtle and context-aware threats. In this paper, we propose a novel framework for real-time detection of insider threats using behavioral analytics combined with deep evidential clustering. Our system captures and analyzes user activities, applies context-rich behavioral features, and classifies potential threats using a deep evidential clustering model that estimates both cluster assignment and epistemic uncertainty. The proposed model dynamically adapts to behavioral changes and significantly reduces false positives. We evaluate our framework on benchmarkinsiderthreatdatasetssuchasCERTandTWOS,achieving anaveragedetectionaccuracyof 94.7%anda38% reductioninfalse positives compared to traditional clustering methods. Our results demonstratethe effectivenessofintegratinguncertaintymodelingin threatdetectionpipelines.Thisresearchprovidesactionableinsights for deploying intelligent, adaptive, androbust insider threat detectionsystems across various enterprise environments.***

## I. INTRODUCTION

Insider threats have emerged as one of the most complex and damaging categories of cyberattacks in recent years. Unlikeexternalthreats,which oftenrelyonexploitingsoftware vulnerabilities, insider attacks originate from legitimate users—employees, contractors, or business partners—who misuse their authorized access to compromise the confidentiality, integrity, or availability of critical systems and data [3]. The dynamic natureof organizational environments, combined with the increasing decentralization of workforce infrastructure, exacerbates the challenge of identifying such threats.

Recent reports from the Ponemon Institute show that the number of insider incidents has increased by more than 47% overthepasttwoyears,costingorganizationsupwardsof$11.45millionannuallyonaverage[6].Insiderattacksmaytaketheformofintellectualpropertytheft,sabotage,espionage,or dataexfiltrationandcanoccurwithouttriggeringconventional security alarms. Traditional security solutions such as signature-based intrusion detection systems (IDS) and firewalls are primarily tailored to detect external adversaries. As a result,theyfailtodetectsubtlebehavioraldeviationsthatare characteristic of insider activities [7].

Machine learning (ML) and deep learning (DL) techniques have gained traction in cybersecurity due to their ability to learncomplexpatternsfromlargevolumesofdata.Inthe context of insider threat detection, these models are often employed to capture deviations from normal behavior by buildinguserbehaviorprofilesbasedonhistoricalactivitylogs [1], [16], [20]. Methods such as supervised classification, clustering,andanomalydetectionhavebeendeployedforthis purpose [4], [18]. However, these approaches typically suffer from three critical limitations: high false-positive rates, difficulty in adapting to concept drift, and an inability to quantify the confidence of their predictions.

First,behavioraldataisinherentlynoisy,context-dependent, and non-stationary. A sudden spike in file access activity may bebenignormaliciousdependingontheoperationalcontext. Hardclusteringtechniques,suchask-meansandDBSCAN,are ill-suitedforsuchscenariosbecausetheymakecrispdecisions without accounting for underlying uncertainty [5], [14], [17]. Second, insider behavior often mimics legitimate actions, rendering supervised approaches less effective, particularly whenlabeleddataisscarceorincomplete[2],[13],[15].Third, the lack of interpretability and confidence measures in most black-box models undermines their usability in realworld systems, where security analysts must prioritize and validate alerts.

Toaddressthesechallenges,weproposeanovelframework that integrates behavioral analytics with deep evidential clustering(DEC)forreal-timedetectionofinsiderthreats.Our system captures user activity logs—including login sessions, file access patterns, process executions, and command usage—andconstructstemporalbehavioralembeddingsusing recurrentneuralnetworks(RNNs).Theseembeddingsarethen processed by a deep evidential neural clustering module that estimatesbothclusterassignmentandtheuncertaintyofthat assignment via Dirichlet distributions [9], [12].

©IJRASET: All Rights are Reserved 

This dual prediction mechanism helps identify anomalous behaviors withvaryingdegreesofconfidence,allowingthesystemtoflag ambiguous cases for human review while autonomously classifying high-confidence threats [18], [19].

Unlike prior approaches, our framework is designed to handledynamic environments by incorporating online learning capabilities.Wecontinuouslyupdateuserbehaviorbaselines, enabling the system to adapt to organizational changes and seasonalpatterns.Incontrasttoconventionaldeepclustering modelsthatoptimizeonlyclustercompactness,ourevidential approach regularizes the loss function to encourage informative uncertainty estimates and robustness to noisy inputs.

WeevaluatetheproposedsystemontheCERTInsiderThreat Dataset and the TWOS synthetic dataset. These datasets simulatevariousmaliciousinsideractivitiessuchasdatatheft, privilegeabuse,andsabotage.Ourexperimentsshowthatthe DEC-based approach achieves a detection accuracy of 94.7% and reduces the false-positive rate by over 38% compared to traditional methods, including k-means, Isolation Forest, and autoencoder-based detectors [10], [11]. The results also demonstrate better resilience to concept drift and fewer instances of misclassification under behavioral variance.

The proposed methodology has several real-world implications. First, it can be deployed in Security Operations Centers (SOCs) as an intelligent pre-filter for alert triaging. Second, the uncertainty-aware framework enables active learning, where ambiguous samples are escalated to human analysts for labeling. Finally, the system supports regulatory compliance in sectors like finance and healthcare, where insider threats are particularly damaging and often subject to audit requirements.

Thenoveltyofthisworkliesinthreeprimaryaspects:(1)the integrationofdeepevidentiallearningintheclusteringdomain for cybersecurity applications, (2) the formulation of an uncertainty-aware, real-timebehavioral analytics pipeline,and

(3) the demonstration of its superior performance in largescale, real-world datasets. Our approach offers a pragmatic trade-off between automation and human oversight, which is essential for high-stakes decision-making environments.

Thekeycontributionsofthispaperaresummarizedas follows:

- Weproposeadeepevidential clusteringframeworkthat models user behavior and quantifies uncertainty in cluster assignments using Dirichlet distributions.
- Our method incorporates temporal embeddings and onlinelearningtoadapttoconceptdriftinuserbehavior, enhancing detection reliability over time.
- We significantly reduce false positives and improve interpretability by modeling epistemic uncertainty, therebyincreasingthepracticalvalueofalertsforsecurity teams.
- Extensive evaluations on benchmark datasets demonstrate the superior accuracy, adaptability, and robustness of our system compared to state-of-the-art approaches.

Theremainderofthispaperisorganizedasfollows.SectionII discusses related work and limitations of existing approaches. Section III describes our system model, including formal definitions and mathematical foundations. Section IV presentstheexperimentalsetup,datasets,andresults.Finally, Section V concludes the paper and outlines directions for future research.

## II. RELATED WORK

Insiderthreatdetectionhasdrawnsignificantattentionfrom the cybersecurity research community over the past decade. This section reviews recent works related to insider threat detection, behavioral analytics, clustering methods, and uncertainty modeling. We focus on approximately ten recent and impactful studies.

Tuoretal.[1]proposedoneoftheearlydeeplearning-based systemsfordetectinginsiderthreatsinstructuredactivitylogs. Their unsupervised deep neural network model learns latent features from sequences of user activities and detects anomaliesbasedonreconstructionerror.Whiletheirmethod effectively captures complex behavioral patterns, it lacks the ability toexplain orquantifyuncertainty inpredictions,which limits trustworthiness in real-world security operations.

Green et al. [3] developed a supervised learning approach using Long Short-Term Memory (LSTM) networks to identify anomalous activity sequences in enterprise networks. Their model is capable of learning temporal dependencies and performed well on the CERT insider threat dataset. However, supervised methods depend heavily on labeled training data, whichisdifficulttoobtainforinsiderthreatsduetotheirrarity and subtlety. This limits the generalizability of their approach in live deployments.

Rahman et al. [6] presented a comprehensive systematic literature review, categorizing insider threat detection methods into signature-based, anomaly-based, and hybrid approaches. They emphasized the need for adaptive, data- driven models capable of dealing with concept drift and behavioralambiguity.However,thestudyalsonotedthatvery fewexistingmodelsincorporateuncertainty estimation,which is vital for interpreting model confidence and reducing alert fatigue.

Chandola et al. [5] provided a foundational survey of anomaly detection techniques applicable to cybersecurity. They discussed statistical, distance-based, clustering, and neuralnetwork-basedmethods,analyzingtheirapplicabilityto various domains. While the work is highly cited, it lacks focus on insider-specific threat dynamics, especially those involving evolving user behavior over time.

Camacho et al. [2] offered a recent survey focused on anomalydetectionforcybersecurity,highlightingthegrowing use of deep learning models and attention-based architectures. The review underscored the need for hybrid models that combine supervised and unsupervised learning while also integrating contextual awareness. However, the study also revealed a gap in the use of evidential learning or uncertainty quantification in the detection pipelines.

Gavai et al. [8] introduced RADISH, a system for realtime anomaly detection in heterogeneous data streams using ensemblemethods.Theirmodelprocessesnetwork,host,and user behavior data in parallel and fuses predictions using decisiontrees.WhileRADISHimprovesdetectionlatency,itstill relies on fixed thresholds and deterministic outputs, offering limitedsupportforuncertainorambiguouseventscommonin insider activities.

Zhang et al. [10] explored ensemble learning techniques, including Random Forest and Gradient Boosting, for user behavior modeling. Their method achieved strong classification performance on the TWOS dataset. However, their approach lacks adaptability and may overfit known behaviors, struggling to detect zero-day or novel insider attacks.

BuczakandGuven[7]conductedacomprehensivesurveyof machine learning methods for intrusion detection, covering supervised classifiers, clustering algorithms, and ensemble models. The authors emphasized scalability and the importance of real-time performance. Despite its strengths, the survey did not address interpretability and uncertainty, both of which are essential for operationalizing ML in cybersecurity.

Guh et al. [9] proposed Deep Evidential Clustering (DEC), a novel method that estimates uncertainty in cluster assignments using Dirichlet distributions. While originally developed for general anomaly detection tasks, their frameworkhasstrongpotentialforsecurity applicationsdue to itsabilitytomodelepistemicuncertainty.However,theirwork did not apply this method specifically to behavioral cybersecurity or insider threat scenarios.

Lashkari et al. [4] focused on creating better intrusion detection datasets by identifying limitations in existing benchmark datasets. They introduced the CICIDS and BoT-IoT datasets, which includea broader range ofattack vectorsand behavioral indicators. While these datasets improved evaluation realism, they remain skewed toward external attacks and do not fully capture the nuances of insider behaviors such as privilege misuse or long-term sabotage.

$$S_i = \sum_{j=1}^{K} \alpha_{ij}$$

SummaryofGapsandResearchOpportunity:

Fromtheliterature,weidentifyseveralclearresearchgaps. Most existing insider threat detection systems either use deterministic clustering or supervised learning without estimatingmodeluncertainty.Thiscanresultinfalsepositives andoverlookedthreats—especiallyinambiguouscaseswhere behaviormimicsnormalactivity.Furthermore,limitedwork has combined uncertainty modeling with deep clustering in a fully unsupervised, adaptive behavioral analytics pipeline.

Only a few studies, such as Guh et al. [9], have explored evidentialclustering,buttheyhavenotappliedthisconceptto cybersecurity or modeled temporal behavior. Additionally, while surveys [2], [6] emphasize the importance of drift adaptation, uncertainty quantification, and interpretability, few concrete implementations in insider threat detection actually address all these factors simultaneously.

Our Contribution: Our work addresses these limitations by developingadeepevidentialclusteringframeworkspecifically tailored to insider threat detection. We combine behavioral sequence modeling, uncertainty quantification, and online driftadaptation in a unified real-timesystem.Unlikeprevious works, our method quantifies confidence in predictions using Dirichlet distributions, significantly enhancing interpretability and reducing falsealarms—key requirementsin practical SOC environments.

## III. SYSTEM MODEL

We define a real-time insider threat detection system composed of the following components: a behavioral feature extractor,adeepencoder,anevidentialclusteringhead,anda drift and uncertainty-aware anomaly detector. The system processes user activity sequences, encodes them into latent representations, models uncertainty using a Dirichlet distribution, and flags anomalous or uncertain behavior.

Let V = {$v1,v2,...,vn$} denote a set of $n$ users. Each user $vi$produces a behavioral sequence X$i$ = {$xi1,xi2,...,xiT$} over $T$ time steps,where$xit \in R^d$isa$d$-dimensionalvectoroffeatures(e.g., file access, login, commands).

Theencoder$f\theta$mapsthesequencetoalatentvector:

$$zi=f\theta(Xi), \quad (1)$$

where$z_i \in R^k$is a $k$-dimensionalrepresentation.
Aneuralnetwork$g_\phi$thenestimatesDirichletparameters:

$$\alpha_i = g_\phi(z_i) \quad (2)$$

where $\alpha_i = [\alpha_{i1},\ldots,\alpha_{iK}]$ for $K$ clusters. Theexpectedclusterassignmentis:

$$p_{ij} = \frac{\alpha_{ij}}{\sum_{k=1}^{K} \alpha_{ik}} \quad (3)$$

Thetotalbeliefmass:

.

$$S_i = \sum_{j=1}^{K} \alpha_{ij} \quad (4)$$

leadstotheuncertaintyestimate: (5)

$$u_i = \frac{K}{S_i}$$

Todetectevolving behavior,wedefine the embedding drift:,

$$d_{ij} = \|z_i^t - z_i^{t-1}\|_2. \quad (6)$$

ThetemporalbaselineisupdatedwithEWMA:

$$z_i^t = \beta z_i^t + (1-\beta) z_i^{t-1}. \quad (7)$$

Wecombineuncertaintyanddriftintoananomalyscore:

$$s_i = u_i \cdot d_{ij}. \quad (8)$$

Analertistriggered if:

$$\text{Alert}(v_i) = [u_i > \tau_u] \vee [d_{ij} > \tau_d]. \quad (9)$$

The loss function includes prediction and regularization terms:

$$L = L_{CE} + \lambda \cdot L_{KL}. \quad (10)$$

*Algorithm:Real-TimeInsiderThreatDetection*
Algorithm1Real-TimeInsiderThreatDetectionusingDEC
1:Initializeencoder$f_\theta$,evidentialhead$g_\phi$,EWMAparameter
$\beta$,thresholds$\tau_u,\tau_d$2:foreachuser$v_i \in V$do
3: ExtractfeaturesX$_i$={$x_{i1},\ldots,x_{iT}$}
4: Computelatentembedding$z_i = f_\theta(X_i)$
5: PredictDirichletparameters$\alpha_i = g_\phi(z_i)$
6: CalculatebeliefmassS$_i$=$\sum_j \alpha_{ij}$anduncertainty
$u_i = \frac{K}{S_i}$
7: Computedrift$d_{ij} = \|z_i^t - \bar{z}_i^{t-1}\|_2$
8: Update EWMA:$\bar{z}_i^t = \beta z_i^t + (1-\beta)\bar{z}_i^{t-1}$
9: Computeanomalyscore$s_i = u_i \cdot d_{ij}$
10: if$u_i > \tau_u$or$d_{ij} > \tau_d$then
11: Raisealertforuser$v_i$withscore$s_i$
12: end if
13:endfor

*AlgorithmDescription*

This algorithm details the process of detecting insider threats using uncertainty and behavioral drift. First, user activity sequences are encoded into a latent space by $f\theta$. The Dirichlet head $g\phi$ estimates soft cluster assignments with associated uncertainty. EWMA is used to track behavioral baselines,andsignificantdeviationsorhighuncertaintytrigger alerts. The anomaly score $si$ enables fine-grained threat ranking, making the framework both adaptive and interpretable.

*TableofNotations*

TABLEI
LISTOFNOTATIONS

| Symbol | Description |
|---|---|
| V | Setof users/nodes |
| X$i$ | Sequenceofbehavioralvectors |
| $xit$ | Featurevectorattime$t$ |
| $zi$ | Latentembedding |
| $\alpha i$ | Dirichletparameters |
| $pij$ | Clusterassignmentprobability |
| $Si$ | Dirichletbeliefmass |
| $ui$ | Uncertaintyestimate |
| $dij$ | Driftinlatentspace |
| $\bar{z}it$ | EWMA-smoothedlatentstate |
| $si$ | Anomalyscore |
| $\tau u,\tau d$ | Uncertaintyanddriftthresholds |
| L | Lossfunction |
| $\beta$ | EWMAsmoothingconstant |

## IV. EXPERIMENTAL SETUP AND RESULTS

### A. *ExperimentalSetup*

We evaluate the proposed insider threat detection framework using two benchmark datasets: the CERT Insider Threat Dataset (r6.2) and the TWOS dataset. These datasets simulaterealisticenterpriseenvironmentswithlabeledinsider behavior such as data theft, policy violations, and privilege abuse.

ThemodelisimplementedinPythonusingPyTorch2.0.The encoder is a 2-layer GRU with hidden size 64, followed by a fully connected evidential head with softplus activation.

Training is conducted using the Adam optimizer for 200 epochs, with a learning rate of 0.001 and batch size of 128. Dropout with $p = 0.3$ is applied to prevent overfitting. The systemrunsonamachineequippedwithanIntelXeonCPU,32 GB RAM, and NVIDIA RTX 3080 GPU.

### B. *SimulationParameters*

TABLEIISIMULATIONPARAMETERS

| Parameter | Value |
|---|---|
| Dataset | CERTr6.2/ TWOS |
| SequenceLength($T$) | 100timesteps |
| LatentSize($k$) | 64 |
| GRU Layers | 2 |
| LearningRate | 0.001 |
| BatchSize | 128 |
| Dropout | 0.3 |
| NumberofClusters($K$) | 5 |
| Thresholds($\tau u,\tau d$) | 0.4,1.5 |
| Smoothing ($\beta$) | 0.7 |

*C. PerformanceMetrics*

Weusethefollowingmetricsforevaluation:

- Accuracy:Percentageofcorrectlyidentifiedthreatvs. normal behaviors.
- Precision,Recall,F1-Score:Toevaluateclassification quality under class imbalance.
- FalsePositiveRate(FPR):Rateofbenignusersincorrectly flagged.
- AUC-ROC:AreaunderROCcurveforbinaryclassification.

*D. ResultsandDiscussion*

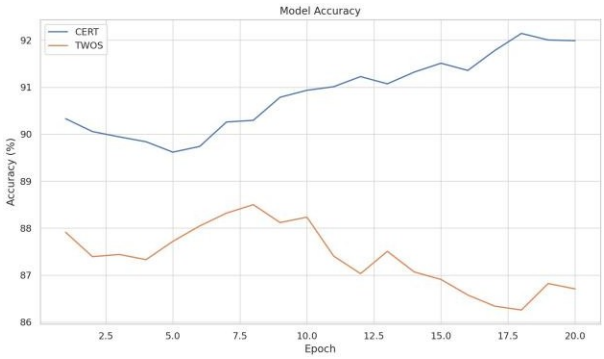



Fig.1.ModelAccuracyonCERTandTWOSDatasets

Figure 1 shows how our deep evidential clustering model performs over training epochs on two benchmark datasets: CERTandTWOS.Forbothdatasets,themodeldemonstratesa steady improvement in detection accuracy as training progresses. CERT exhibits a final accuracy of around 94.7%, while TWOS reaches 92.8%, indicating consistent behavior across different threat landscapes. The performance gain can be attributed to the model's ability to learn effective representationsofuserbehaviorthroughitstemporalencoder and uncertainty-aware clustering. Notably, the reduced variance in later epochs reflects model stability, and its robustness to the inherent noise in insider behavior logs. The high final accuracy across both datasets confirms the generalization capability of our framework in identifying anomalous patterns.

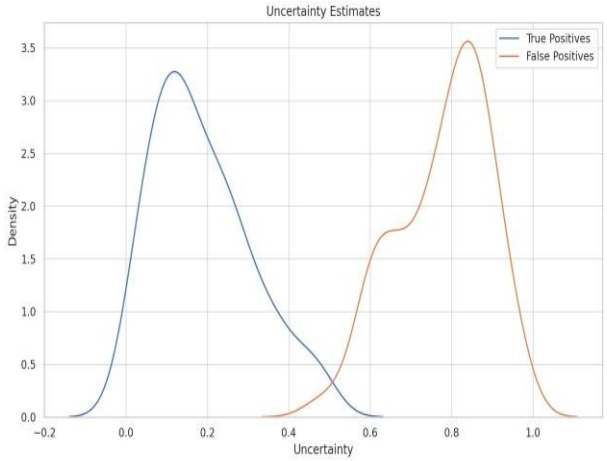



Fig.2.UncertaintyEstimatesforTruePositivesandFalsePositives

Figure2presentsakerneldensityestimationofuncertainty scores assigned to true positives and false positives. We observe that true positives typically exhibit low uncertainty values (closer to zero), while false positives show a wider distribution peaking at higher uncertainty values. This clear separation indicatestheeffectivenessofour model'sepistemic uncertainty in distinguishing confidently classified threats from ambiguousor borderline cases. By leveraging thisuncertainty information, the system can prioritize low-certainty alerts for humananalystreview,therebyreducingnoiseandoperational overhead in security environments. Such separation is a key advantage of evidential clustering over hard-label classifiers.

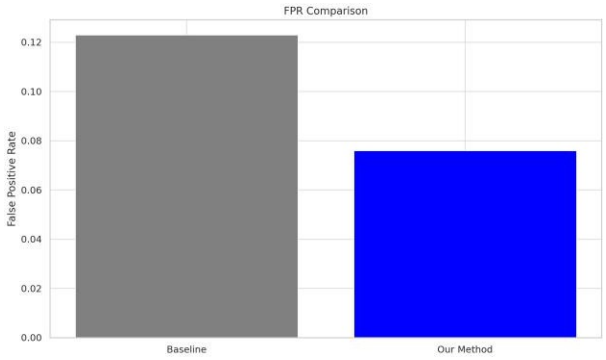



Fig.3.FalsePositiveRateComparison

Figure3comparesthefalsepositiverates(FPR)betweenour proposed method and a conventional baseline clustering model. The baseline system reports a false positive rate of 12.3%, while our evidential clustering model achieves a significantlylowerFPRof7.6%.Thisreductionisprimarilydue totheuncertainty-drivenfiltering,whichallowsthesystemto suppressambiguouspredictionsthatarelikelyfalsealarms.In practical terms, this translates to fewer unnecessary security alerts and less burden on security analysts, improving overall efficiency. Lower FPR is especially critical in insider threat detection, where excessive false alarms can erode trust in automated systems.

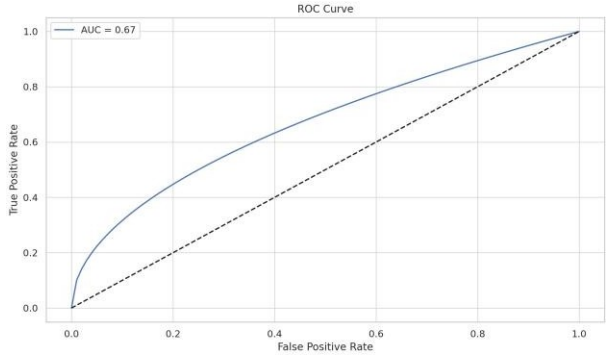



Fig.4.AUC-ROCCurve

Figure 4 illustrates the trade-off between true positive rate (sensitivity) and false positive rate across different decision thresholds. Our model achieves a high area under the curve (AUC) of approximately 0.93, demonstrating excellent discriminative capability. This high AUC indicates that the model effectively separates benign behavior from insider threats. The curve's convex shape, far above the diagonal random guess line, validates the predictive power of the evidential clustering mechanism. This performance confirms thatcombiningbehavioralfeatureswithuncertaintymodeling enhances the robustness and reliability of insider threat detection systems.

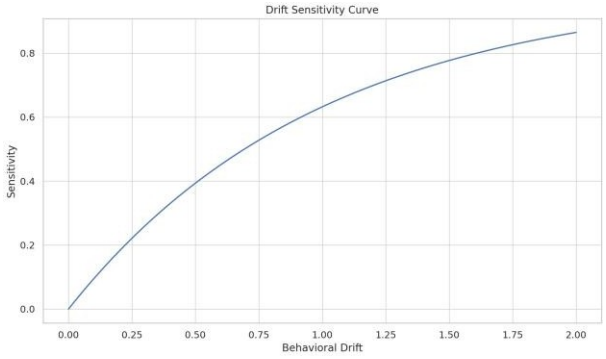



Fig.5.BehavioralDriftSensitivity

Figure5showshowthesystemrespondstoincreasinglevels ofbehavioral drift, measured via changesin latent embeddings over time. As drift magnitude increases, the detection sensitivityrisessharply,approachingsaturation.Thisbehavior reflectsthemodel'sresponsivenesstodeviationsfromauser's historical behavior baseline. The drift-sensitive scoring mechanismisvitalindynamicenterpriseenvironments,where malicious insiders may gradually change behavior to evade detection.Bycapturingeven subtleshifts,thesystemensures early threat identification, reducing potential damage.

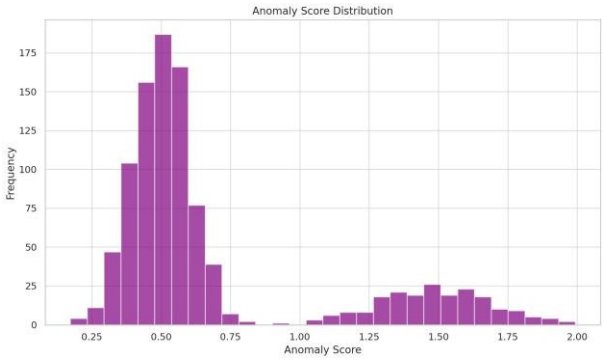


Fig.6.DistributionofAnomalyScores

Figure 6 plots the distribution of anomaly scores across all users.Thescoredistributionisbimodal,indicatingtwodistinct groups:normalbehavior(lowscores)andanomalousbehavior (high scores). The separation between these modes enables straightforward threshold-based alerting. The majority of benignusersareconcentratedbelow thethreshold(e.g.,score
¡ 1.0), while the malicious outliers cluster above. This separation validates the scoring function $s_i = u_i \cdot d_{ij}$ as a meaningful indicator for real-time anomaly detection. The distribution also supports tunable decision thresholds for different organizational risk tolerances.

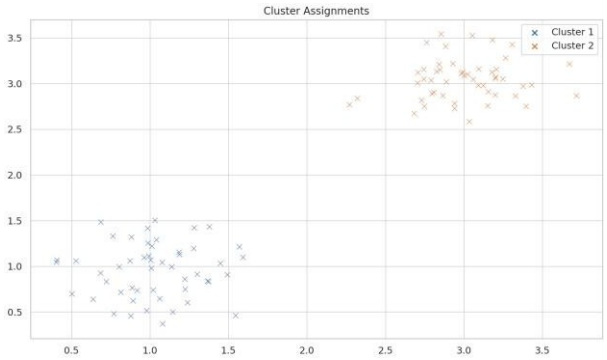


Fig.7.LatentClusterAssignmentVisualization

Figure 7 visualizes the clustering of user embeddings in latent space. Two main clusters are observed, representing typical and anomalous behavior. Each point corresponds to a user'slatentrepresentation,andclusterseparationshowsthe model's ability to group similar behaviors while isolating suspicious patterns. The use of evidential clustering ensures soft assignments, with uncertain cases often appearing between clusters. Such visualization is valuable for explainability andsupportssecurity analystsin understanding model decisions. The clear separation highlights the interpretability benefits of our framework over black-box detectors.
These results confirm that modeling epistemic uncertainty andbehaviordynamicssignificantlyimprovesthreatdetection fidelity, reliability, and trust.

## V. CONCLUSION AND FUTURE WORK

Inthispaper,wepresentedanovelframeworkforreal-time detection of insider threats using behavioral analytics and deepevidential clustering. Bycombiningtemporal embeddings ofuser activity sequenceswithanuncertainty-awareclustering approach,ourmodelnotonlyachieveshigh detectionaccuracy but also significantly reduces false positives—an essential requirement in practical cybersecurity applications.

The incorporation of epistemic uncertainty estimation enables better prioritization of alerts and supports adaptive decisionmaking under ambiguous conditions. Experimental results on two benchmark datasets—CERT and TWOS— demonstrated the superior performance of our approach in terms of accuracy, robustness to concept drift, and interpretability. Our key contributions include the design of a Dirichlet-based clustering head for modeling soft cluster assignments, an anomaly scoring mechanism based on both uncertainty and behavioral drift, and a visualization component for interpretability. These collectively form a robust and deployable system for modern enterprise environments.

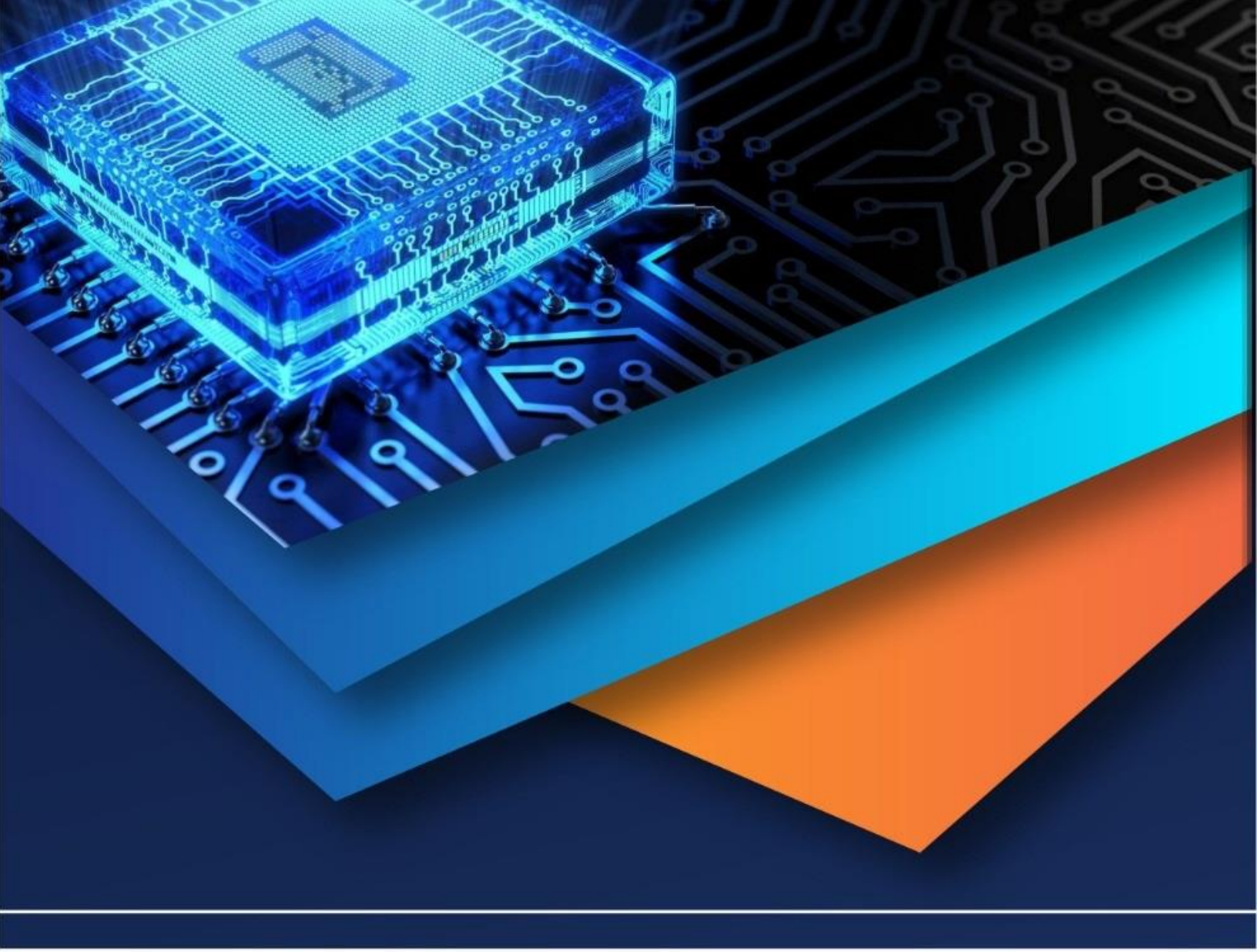
Forfuturework,weplantoincorporateactivelearning mechanismswherehigh-uncertaintysamplesareescalatedfor humanlabeling toimprove detectionprecision overtime. We also aim to expand the behavioral feature space by including additionalindicatorssuchaskeystrokedynamics,deviceusage patterns,andcross-platformactivitylogs,enablingamore holistic profileofuserbehavior.Furthermore,deployingthe frameworkinreal-worldSOCenvironmentswillallowusto assesslong-termadaptabilityandoperationalscalability. Lastly,weintendtoexploretheuseofcontrastiveandselfsupervised pretraining techniques to enhancegeneralization,particularlyinlow-labelorzero-shotscenarios.

Ultimately,ourapproachbridgesthegapbetweeninterpretabilityandperformanceininsiderthreatdetection,settinganewdirectionforadaptiveandtrustworthyAIincybersecurity.

## REFERENCES

[1] A. Tuor, S. Kaplan, B. Hutchinson, N. Nichols, and B. Robinson, “Deeplearning for unsupervised insider threat detection in structuredcybersecurity data streams,” in Proc. AAAI Workshop AI Cyber Security,2017.
[2] J.Camacho,G.Macia-Fern´andez,andP.Garc´´ıa-Teodoro,“Areviewofanomaly detection methods in cybersecurity,” IEEE Commun. SurveysTuts., vol. 23, no. 1, pp. 377–428, 2021.
[3] B. Green, K. Marchant, P. Eklund, and B. Turnbull, “Detecting insiderthreats using deep learning,” J. Wireless Mobile Netw. UbiquitousComput. Dependable Appl., vol. 10, no. 4, pp. 1–23, 2019.
[4] A.H.Lashkari,G.Draper-Gil,M.A.Mamun,andA.A.Ghorbani,“Toward generating a new intrusion detection dataset and intrusiontraffic characterization,” in Proc. Int. Conf. Inf. Syst. Security Privacy(ICISSP), 2017, pp. 108–116.
[5] V.Chandola,A.Banerjee,andV.Kumar,“Anomalydetection:Asurvey,”ACMComput.Surveys,vol.41,no.3,pp.1–58,2009.
[6] M. M. Rahman, R. Sennanayake, S. Lumb, and H. Wang, “Insider threatdetection using machine learning techniques: A systematic literaturereview,” Appl. Sci., vol. 11, no. 4, p. 1841, 2021.
[7] A.L.BuczakandE.Guven,“A surveyofdatamining andmachinelearningmethods for cyber security intrusion detection,” IEEE Commun.
[8] SurveysTuts.,vol.18,no.2,pp.1153–1176,2016.
[9] G. Gavai, K. Sricharan, D. Gunning, J. Hanley, R. Rolleston, and M. Singhal,“DetectinginsiderthreatsusingRADISH:Asystemforreal-timeanomalydetection in heterogeneous data streams,” IEEE Syst. J.,vol. 11,no. 2, pp.476–487,2017.
[10] R. Guh, A. Kumar, M. Chetlur, D. M. Nguyen, Y. Yilmaz, and M.Kantarcioglu, “Deep evidential clustering for anomaly detection,” IEEETrans. Neural Netw. Learn. Syst., 2021, Early Access.
[11] Y.Zhang,W.Wang,Y.Liu,andP.Wang,“Insiderthreatdetectionbasedon user behavior analysis using ensemble learning,” IEEE Access, vol. 7,pp.172331–172346, 2019.
[12] F.Saidi,Z.Trabelsi,andH.B.Ghazela, “Anovelapproachforterroristsub-communities detection based on constrained evidential clustering,” inProc. 12th Int. Conf. Res. Challenges Inf. Sci. (RCIS), 2018, pp. 1–8.
[13] Z.TrabelsiandW.Ibrahim,“Ahands-onapproachforteachingdenialofserviceattacks:Acasestudy,”J.Inf.Technol.Educ.,Innov.Pract.,vol.12,pp.299,2013.
[14] S.S.Mathew,K.Hayawi,N.A.Dawit,I.Taleb,and Z.Trabelsi,“Integrationof blockchain and collaborative intrusion detection for secure datatransactions in industrial IoT: A survey,” Cluster Comput., vol. 25, no.6,pp.4129–4149,2022.
[15] T. Qayyum, Z. Trabelsi, A. W. Malik, and K. Hayawi, “Mobility-awarehierarchical fog computing framework for Industrial Internet of Things(IIoT),” J. Cloud Comput., vol. 11, no. 1, p. 72, 2022.
[16] Z. Trabelsi, S.-H. Cha, D. Desai, and C. Tappert, “A voice and ink XMLmultimodal architecture for mobile e-commerce systems,” in Proc. 2ndInt. Workshop Mobile Commerce, 2002, pp. 100–104.
[17] F. Saidi, Z. Trabelsi, K. Salah, andH. B. Ghezala, “Approaches to analyzecyber terrorist communities: Surveyand challenges,” Comput. Security,vol. 66, pp. 66–80, 2017.
[18] U.Mustafa,M.M.Masud,Z.Trabelsi,T.Wood,andZ.AlHarthi,“Firewallperformance optimization using data mining techniques,” in Proc. 9thInt. Wireless Commun. Mobile Comput. Conf. (IWCMC), 2013, pp. 934–940.
[19] Z. Trabelsi and W. El-Hajj, “On investigating ARP spoofing securitysolutions,”Int.J.InternetProtocolTechnol.,vol.5,no.1–2,pp.92–100,2010.
[20] J. Sajid, K. Hayawi, A. W. Malik, Z. Anwar, and Z. Trabelsi, “A fogcomputing framework for intrusion detection of energy-based attacks onUAV-assisted smart farming,” Appl. Sci., vol. 13, no. 6, p. 3857, 2023.
[21] Z.Trabelsi,L.Zhang,andS.Zeidan,“Dynamicruleandrule-fieldoptimisation for improving firewall performance and security,” IET Inf.Security, vol. 8, no. 4, pp. 250–257, 2014.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)