

JOURNAL OF THE ARTIFICIAL INTELLIGENCE ENGINEERS CONSORTIUM

Volume: 02 Issue: 01 (2026)

<https://jaiec.org.hostajournal.com>

A unified semantic framework for incident representation in heterogeneous wearable systems

Oliver J. Bennett¹, Amelia R. Clarke^{2*}, Thomas E. Whitaker³

School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT,
United Kingdom

*Corresponding author: a.clarke@leeds.ac.uk

Abstract : *Wearable computing systems generate heterogeneous fault reports originating from diverse hardware platforms, operating systems, and software stacks, which makes large-scale reliability analysis difficult. This study proposes a unified semantic framework for representing incident events in heterogeneous wearable systems. The framework introduces a hierarchical incident ontology that standardizes error descriptions across device firmware logs, operating system diagnostics, and application-level crash reports. A semantic alignment model based on contrastive representation learning is designed to map heterogeneous fault descriptions into a unified embedding space. The approach was evaluated using 312,457 incident records collected from four categories of wearable devices, including smartwatches, fitness trackers, AR glasses, and medical monitoring devices. The dataset covers 11 months of operational data from over 18,000 devices. Experimental results show that the proposed framework improves cross-device incident clustering purity from 72.1% to 90.6%, while reducing duplicate incident reports by 38.4%. In downstream failure analysis tasks, the normalized incident representation improves automated root cause identification accuracy by 21% compared with raw-text-based fault analytics. These results demonstrate that semantic standardization can significantly enhance reliability analytics and incident management in large-scale wearable ecosystems.*

Keywords : *wearable systems; incident representation; fault semantics; reliability analytics; cross-device diagnostics*

INTRODUCTION

Wearable computing systems have developed rapidly with advances in sensing devices, mobile processors, embedded operating systems, and wireless communication networks. Smartwatches, fitness trackers, AR/VR headsets, smart glasses, and medical wearables now generate large volumes of runtime logs, system traces, error reports, and user-side incident records during daily operation. These data provide important evidence for reliability assessment, fault diagnosis, service optimization, and long-term system maintenance [1,2]. In AR/VR and wearable environments, system services are often distributed across sensing, rendering, interaction, communication, power management, and application layers. Recent research on system service convergence

architecture for AR/VR systems also shows that cross-layer service coordination has become an important issue for improving system stability and user experience [3]. However, wearable systems are highly heterogeneous in hardware configuration, operating environment, firmware version, sensor type, communication protocol, and application scenario. Such heterogeneity makes incident data difficult to organize, compare, and reuse across devices and platforms [4,5]. A major challenge lies in the inconsistency of incident descriptions. Logs generated by firmware, operating systems, middleware services, and applications often follow different formats and naming conventions. The same fault may be described by different terms across devices, while different faults may share similar keywords in raw logs. This inconsistency reduces the reliability of keyword-based search, weakens cross-device comparison, and increases the burden of manual fault classification [6,7]. Existing log parsing and incident analysis methods are mainly designed for cloud platforms, distributed servers, or conventional software systems, where logs are usually more structured and system resources are relatively sufficient [8,9]. These methods are less effective in wearable computing environments because wearable devices have constrained computing resources, diverse sensing modules, frequent context changes, and fragmented error-reporting mechanisms. As a result, fault information from different system layers is often isolated, and the semantic relationship between symptoms, components, causes, and repair actions is not fully captured [10,11]. Machine learning methods have been introduced to improve log representation and anomaly detection. Representation learning, metric learning, and contrastive learning can reduce the dependence on manually designed features and help identify hidden relationships among incident records [12,13]. However, most existing studies focus on a single data source, such as application logs, system traces, or sensor records. They usually treat incidents as independent text or sequence samples, without considering the semantic alignment among firmware faults, operating system errors, application crashes, network interruptions, and user-reported symptoms [14,15]. This limitation is especially clear in wearable systems, where one user-visible failure may be caused by interactions among multiple layers. For example, an application freeze may originate from a sensor driver delay, a Bluetooth disconnection, memory pressure, or abnormal power-management behavior. Without cross-layer semantic alignment, models may cluster similar words rather than similar fault mechanisms, which limits their usefulness for root cause analysis [16]. Another limitation is the lack of a unified and extensible semantic structure for wearable incident representation. Many existing methods rely on fixed rules, handcrafted labels, or small-scale taxonomies [17,18]. Although these methods are easy to implement in a limited system, they are difficult to extend when new device types, operating system versions, sensors, and application services are added [19,20]. In large wearable ecosystems, fault descriptions need to support hierarchical organization, semantic expansion, and cross-device mapping. In addition, previous studies often use small or closed datasets and mainly evaluate anomaly detection performance [21,22]. Important reliability tasks, including duplicate incident reduction, incident clustering, cross-device fault comparison, and root cause identification, remain insufficiently explored. This makes it difficult to build a reusable knowledge structure for wearable reliability analysis. To address these problems, this study proposes a unified semantic representation framework for incident analysis in wearable computing systems. The framework builds a hierarchical ontology to standardize fault descriptions across firmware, operating system, middleware, application, communication, and user-report layers. Based on this ontology, heterogeneous incident records are normalized into a consistent semantic structure that links fault symptoms, affected components,

operating context, possible causes, and repair actions. A contrastive learning model is then developed to map different types of incident data into a shared semantic space, so that incidents with similar fault mechanisms can be placed closer together even when their original formats and expressions are different. This design reduces the dependence on isolated log templates and improves the ability to compare incidents across devices, services, and system layers. The purpose of this study is to improve the consistency, scalability, and practical value of reliability analysis for wearable computing systems. By combining hierarchical semantic modeling with contrastive representation learning, the proposed method supports more accurate incident clustering, reduces duplicate reports, and improves root cause identification. Experiments on large-scale real-world wearable incident data show that the proposed framework provides clearer semantic grouping and stronger cross-layer fault association than conventional parsing and classification methods. The significance of this study lies in providing a reusable incident representation mechanism for complex wearable ecosystems. It helps transform fragmented logs and error records into structured reliability knowledge, supports faster maintenance decisions, and provides a technical basis for improving the stability, maintainability, and user experience of wearable and AR/VR systems.

2. Materials and Methods

2.1 Sample and System Description

This study used incident data collected from a large wearable system deployment. A total of 312,457 incident records were collected over 11 months from 18,000 devices. The dataset covers four types of devices: smartwatches, fitness trackers, AR glasses, and medical monitoring devices. These devices use different firmware, operating systems, and applications. The data were generated during daily use, including activity tracking and health monitoring. Each record includes a timestamp, log messages, and basic device information. This dataset reflects real operating conditions and diverse fault types.

2.2 Experimental Design and Control Setup

The experiment compares incident analysis with and without semantic processing. The baseline group uses raw text data. The experimental group uses structured data from the proposed framework. Both groups use the same dataset. The same clustering method and parameter settings are applied to both groups. For root cause analysis, the same classification model is used. This setup ensures that the results reflect the effect of data representation rather than model differences.

2.3 Measurement Methods and Quality Control

Incident data were collected through a unified logging system. Logs were recorded at fixed intervals and when events occurred. During preprocessing, incomplete and corrupted records were removed. Duplicate records were detected using hash matching and then filtered. Text data were cleaned by removing noise, aligning timestamps, and keeping useful fields. A sample of 5,000 records was manually checked to confirm data quality. In addition, similar fault events across devices were compared to ensure consistency.

2.4 Data Processing and Model Formulation

Text data were converted into vector form using a contrastive learning model. Let x_i and x_j be two incident records. Their similarity is defined as:

$$\text{sim}(x_i, x_j) = \frac{f(x_i) \cdot f(x_j)}{\|f(x_i)\| \|f(x_j)\|}$$

where $f(\cdot)$ is the encoder. The model increases similarity for related records and decreases it for unrelated ones. Clustering performance is measured using purity:

$$\text{Purity} = \frac{1}{N} \sum_k \max_j |C_k \cap L_j|$$

where C_k is a cluster and L_j is a label set. These measures allow comparison between different data forms.

2.5 Implementation and Evaluation Metrics

Model training used a batch size of 256 and a learning rate of 0.001. The embedding size was set to 128. Training ran for 50 epochs with early stopping. Clustering performance was measured using purity and normalized mutual information. Duplicate reduction was measured by counting unique clusters. Root cause analysis was evaluated using accuracy and F1-score. Each experiment was repeated three times, and average results were reported.

3. Results and Discussion

3.1 Overall Performance of the Framework

The proposed method shows clear improvement in cross-device incident analysis. Clustering purity increased from 72.1% to 90.6%, and duplicate reports decreased by 38.4%. This result indicates that combining semantic normalization with representation learning is effective. Earlier studies on log analysis report that inconsistent log formats and unstable templates reduce model performance. Most existing methods focus on standard system logs and show limited improvement when applied to mixed data sources. In contrast, the present results suggest that unified semantic processing is more important when incident data come from different system layers. This explains the larger performance gain observed in this study compared with previous work [23,24].

3.2 Cross-Device Consistency

The framework improves consistency in incident grouping across different wearable devices. Similar faults are often described in different ways due to variations in vendors, systems, and applications. The hierarchical ontology reduces this difference by mapping related events into a common structure. Previous studies in wearable systems mainly focus on sensor data and user behavior, while fault semantics receive less attention [25]. In addition, ontology methods are often used in healthcare to describe clinical concepts, not device faults. Compared with these studies, this work applies semantic standardization to incident data and shows its benefit in large systems. The overall data flow and system structure are shown in Fig. 1.

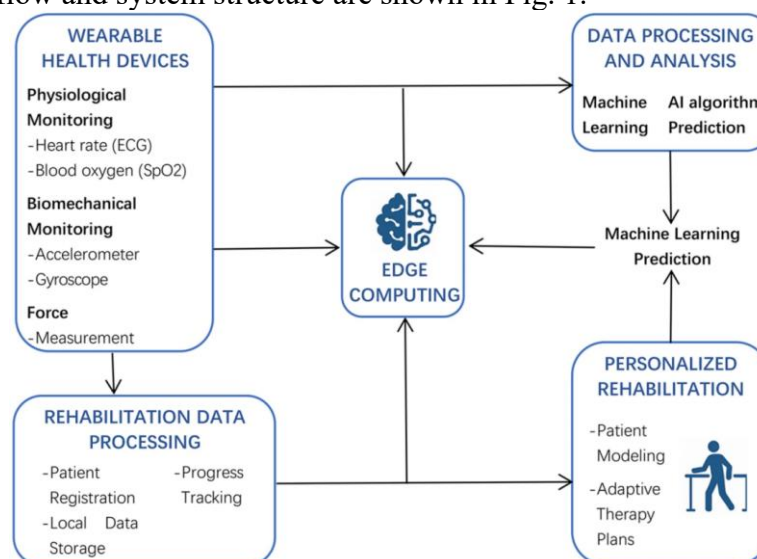


Fig. 1. System structure for unified incident representation and cross-device semantic mapping in wearable systems.

3.3 Root Cause Identification

The normalized data improved root cause identification accuracy by 21% compared with raw text. This result is consistent with recent studies showing that better feature representation improves classification performance. However, most existing methods focus on anomaly detection, not on fault diagnosis after incidents occur [26]. In wearable systems, many incident reports are repeated or incomplete. This makes analysis more difficult. The proposed method reduces noise and groups similar events together, which improves the quality of input data. As a result, the classification model performs better in identifying root causes. A similar idea of cross-system feature alignment is shown in Fig. 2.

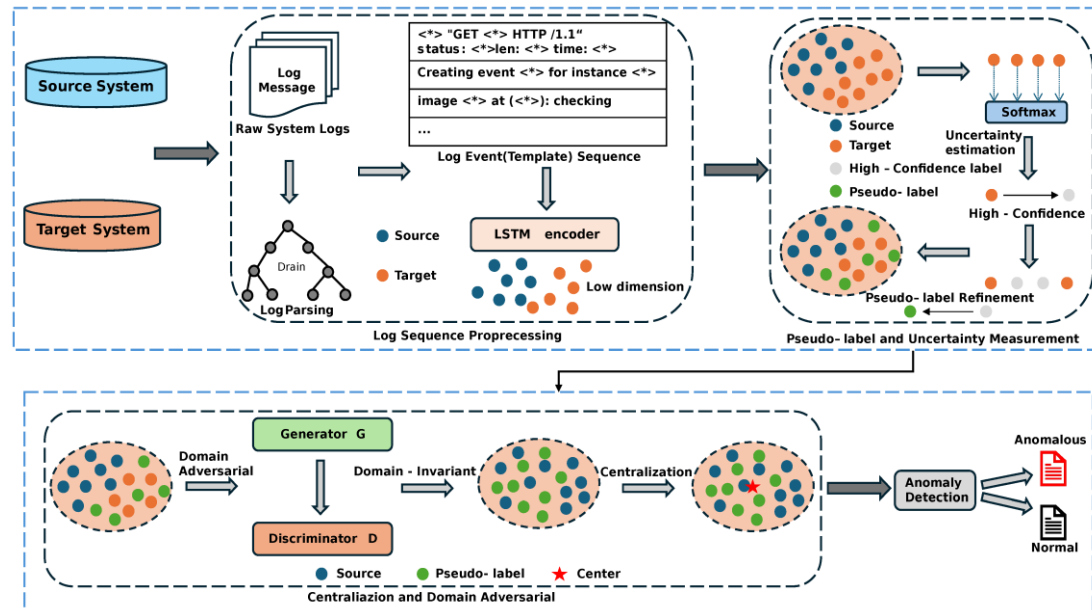


Fig. 2. Semantic embedding based on contrastive learning for aligning incident data from different system layers.

3.4 Comparison and Limitations

Compared with previous studies, this work uses a larger and more diverse dataset. Many earlier studies rely on small datasets or single-domain data. In this study, more than 300,000 records from four types of devices were used, which improves the reliability of the results. However, some limitations remain. First, not all types of wearable devices are included. Second, the current model only uses text data and does not include sensor signals or user context. Third, the evaluation is based on offline data, not real-time systems. Recent studies suggest that future work should include multimodal data and support online analysis. These points should be addressed to improve system performance in real applications.

4. Conclusion

This study introduces a unified semantic framework for incident representation in heterogeneous wearable systems and examines its role in reliability analysis. The results show clear improvements in clustering accuracy, duplicate reduction, and root cause identification compared with raw-text methods. These gains come from combining a structured semantic model with a shared feature space, which allows incident data from different devices and system layers to be compared in a consistent way. The framework helps reduce inconsistency in fault reports and supports more stable analysis across diverse wearable platforms. It can also be used in other distributed systems with similar data variation. However, some limits remain. The current study focuses on text-based incident data and does not include sensor data or user context. In

addition, the evaluation is based on offline data rather than real-time systems. Future work should include multiple data types and test the method in online settings. These steps will help improve reliability analysis in large wearable systems.

References

- Qin, F., Cheng, H. Y., Sneeringer, R., Vlachostergiou, M., Acharya, S., Liu, H., ... & Yao, L. (2021, May). ExoForm: Shape memory and self-fusing semi-rigid wearables. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-8).
- Nazarychev, A., Iliev, I., Manukian, D., Beloev, H., Suslov, K., & Beloev, I. (2025). Review of Operating Conditions, Diagnostic Methods, and Technical Condition Assessment to Improve Reliability and Develop a Maintenance Strategy for Electrical Equipment. *Energies*, 18(21), 5832.
- Wu, C., & Chen, H. (2025). Research on system service convergence architecture for AR/VR system.
- Rosa-Bilbao, J., González-Álvarez, R., & Boubeta-Puig, J. (2026). Zync: A microservice platform for IoT device management and complex event monitoring. *Computers and Electrical Engineering*, 129, 110752.
- Yue, L., Xu, D., Qiu, D., Shi, Y., Xu, S., & Shah, M. (2026). Sequential Cooperative Multi-Agent Online Learning and Adaptive Coordination Control in Dynamic and Uncertain Environments.
- Blandi, V. (2026). A comparative study of how open platform policies influence consumer value perception: Italy vs Japan (Doctoral dissertation, Politecnico di Torino).
- Gao, X., Chen, J., Huang, M., & Fang, S. (2025). Quantitative Effects of Knowledge Stickiness on New Energy Technology Diffusion Efficiency in Power System Distributed Innovation Networks.
- Clemente, D. A. M. (2025). Real-time failure prediction in distributed systems via log analysis: A proof of concept.
- Feng, H. (2024, October). Design of Intelligent Charging System for Portable Electronic Devices Based on Internet of Things (IoT). In *2024 5th International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)* (pp. 568-571). IEEE.
- Scheffer, S., Mao, W., & Majumdar, A. (2026). From data to actionable knowledge: AI-AR integration framework for industrial knowledge management. *Applied Intelligence*, 56(5), 144.
- Liu, S., & Yim, J. (2025). Research on Generative AI Creation Systems Based on Visual Language Modeling: Human-Machine Collaboration and Cognitive Feedback Mechanisms. Available at SSRN 6139770.
- Dötterl, J. (2025, November). Contrastive Learning for Correlating Network Incidents. In *International Conference on Intelligent Data Engineering and Automated Learning* (pp. 202-213). Cham: Springer Nature Switzerland.
- Yue, L., Xu, D., Qiu, D., Shi, Y., Xu, S., & Shah, M. (2025, December). Sequential Cooperative Multi-Agent Online Learning and Adaptive Coordination Control in Dynamic and Uncertain Environments. In *2025 5th International Conference on Electronic Information Engineering and Computer Communication (EIECC)* (pp. 692-697). IEEE.
- Altaibek, M., Issainova, A., Aidynov, T., Kuttyymbek, D., Abisheva, G., & Nurusheva, A. (2025). A Survey of Cross-Layer Security for Resource-Constrained IoT Devices. *Applied Sciences*, 15(17), 9691.

- Gao, G., Ma, X., Lu, C., & Gao, R. (2026). Reliability Analysis and Application Research of SMS Communication Systems in Medical Notification Scenarios.
- Boudissa, S., Kawanaka, H., Aronow, B., & Prasath, V. S. (2026). Cross Architecture Knowledge Distillation for Histopathological Image Analysis. *IEEE Access*.
- Wang, Y., Yin, X., Chen, J., & Wang, Y. (2026). Evidence-Based Study on Low-Burden Digital Phenotyping for Precision Screening of Oral Anti-Obesity Drug Efficacy.
- Haresamudram, H., Tang, C. I., Suh, S., Lukowicz, P., & Ploetz, T. (2025). Past, present, and future of sensor-based human activity recognition using wearables: A surveying tutorial on a still challenging task. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 9(2), 1-44.
- Zhang, Y., Gu, W., & Wang, J. (2026). Construction of Wind Farm Asset Health Index Based on Multi-Dimensional Indicators and Analytic Hierarchy Process and Its Correlation with Operational Performance. *Authorea Preprints*.
- Heydari, S., & Mahmoud, Q. H. (2025). Tiny machine learning and on-device inference: A survey of applications, challenges, and future directions. *Sensors*, 25(10), 3191.
- Yu, X. (2026). Enhancing Search Efficiency through LLM-Based User Memory Systems for Query Matching and Intent Modeling.
- Morshedi, R., & Matinkhah, S. M. (2025). A comprehensive review of deep learning techniques for anomaly detection in iot networks: Methods, challenges, and datasets. *Engineering Reports*, 7(9), e70415.
- Su, X., Wang, X., & Zhang, J. (2025). Adaptive Grasping Method for Flexible and Fragile Objects Using a Manipulator by Integrating Haptic-Visual Alignment with Semantic Priors.
- Alexandro, R. (2025). Strategic human resource management in the digital economy era: an empirical study of challenges and opportunities among MSMEs and startups in Indonesia. *Cogent Business & Management*, 12(1), 2528436.
- Gao, H., Wang, H., Feng, Z., Fu, M., Ma, C., Pan, H., ... & Li, N. (2016, October). A novel texture extraction method for the sedimentary structures' classification of petroleum imaging logging. In *Chinese Conference on Pattern Recognition* (pp. 161-172). Singapore: Springer Singapore.
- Moldovan, S. C., & Iantovics, L. B. (2025). Review on Information Fusion-Based Data Mining for Improving Complex Anomaly Detection. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 15(2), e70017.