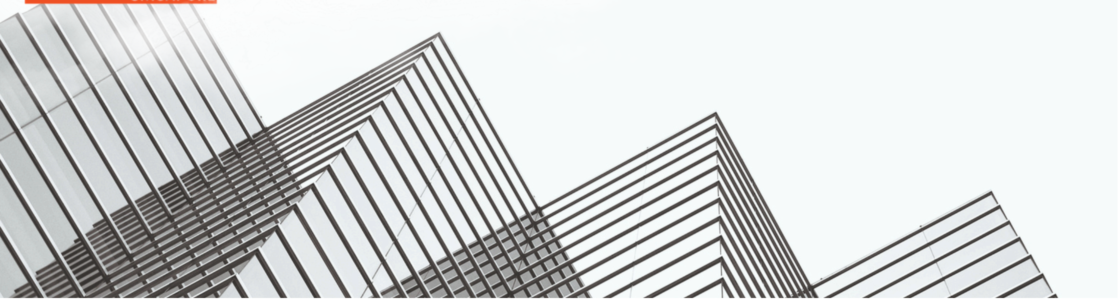


THRUST 4: EVALUATION

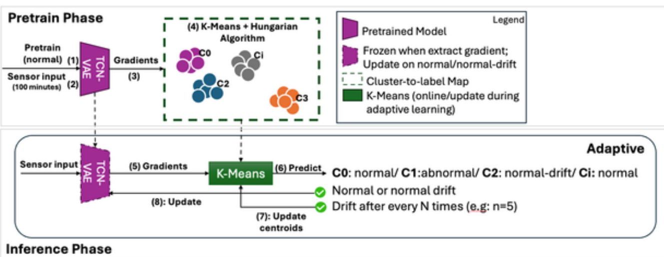
DRIFT-AWARE TEMPORAL CONVOLUTIONAL NETWORKS (TCN)-VARIATIONAL AUTOENCODERS (VAE)

A primary challenge in detecting critical infrastructure anomalies is that modern cyber-physical systems (CPS) operate under dynamic conditions, including occupancy patterns, operational settings, and environmental conditions that change over time. As system behavior evolves, such changes may be misclassified as anomalies in traditional detection models, leading to false alarms. Another potential risk is with adaptive models. They inadvertently learn from attack data, resulting in model contamination and degraded detection accuracy.
The drift-aware TCN-VAE framework is designed to address these challenges. It combines both Temporal Convolutional Network (TCN) and Variational Autoencoder (VAE) that utilizes input gradients as behavioral fingerprints.

- Temporal Convolutional Network (TCN): Captures temporal patterns and sequences
- Variational Autoencoder (VAE): Learns robust representations
- Gradient Profiles: Distinguish between behavior types by measuring the sensitivity reconstruction loss to each input feature

To evaluate the effectiveness of this approach, a case study is conducted using a real-world Building Management System (BMS) dataset, with a focus on temperature setpoint attacks on HVAC systems.

Under normal operations, the sensor dynamics are stable and highly correlated due to feedback control. Benign changes, such as after-office hours, maintenance activities, and gradual control retuning, result in structured but shifted behaviors which are known as “normal drifts.” In contrast, attacks introduce abrupt and inconsistent patterns that break these normal correlations.



The system processes data streams sequentially to extract gradient profiles. These profiles serve as behavioral fingerprints, enabling the classification of data points as Normal, Normal Drift, or Abnormal. Upon detecting normal and normal drift samples, the framework incrementally updates the model to stay current with evolving system behavior. After a configurable number of normal drift batches, the system updates the K-Means centroid to balance adaptability with stability. This ensures that the model remains robust under long-term system evolution while preventing contaminated learning from faulty behaviors, misconfigurations, and malicious activities.

Related publication: Le, T. L., Tan, H. C., Ng, Z. W., Kalbarczyk, Z., Yau, D. K., Mashima, D., & Lou, X. (2026, June). Drift-Aware Online Anomaly Detection in Smart Buildings via Temporal Variational Autoencoder Gradient Profile. In Proceedings of the 17th ACM International Conference on Future and Sustainable Energy Systems (pp. 272-283).