Journal of Civil Engineering and Urban Planning (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/jceup.2026.080202
ISSN 2616-3969 Vol. 8 Num. 2

A Real-Time Reliability Assessment Method for Urban Bridge Erecting Machines Based on a Three-State Noisy-OR Bayesian Network and XGBoost Evidence Fusion

Rui Guo

China Academy of Machinery Science and Technology Group, Beijing, China

Keywords: Urban bridge erecting machine; Bayesian network; three-state Noisy-OR; XGBoost; Real-time reliability assessment

Abstract: The hoisting trolley system of an urban bridge erecting machine is a strongly coupled electro-mechanical-hydraulic subsystem with very few field samples for its critical components. In real-time reliability assessment, two issues are normally treated separately: low-level multi-source condition monitoring and system-level probabilistic reasoning. To bridge this gap, this paper proposes a real-time reliability assessment method based on a three-state Noisy-OR Bayesian network (BN) coupled with XGBoost evidence fusion. First, the state space of the bottom-layer Bayesian network nodes is extended from the conventional Normal/Failed binary form to a Normal/Degraded/ Failed three-state form, with a hybrid modelling strategy in which mechanical/hydraulic components are described by three states and electrical/logical components remain binary. Second, a simplified three-state Noisy-OR risk-propagation rule is adopted to aggregate the Normal, Degraded and Failed probabilities of bottom-layer nodes upward into system-level state probabilities. Finally, XGBoost is used to convert multi-source monitoring features into an initial SoftMax three-class probability vector; after soft-evidence stabilization, the corresponding prototype evidence vector is injected into the leaf nodes of the Bayesian network, closing the loop from data perception to topological reasoning. Taking a hoisting-trolley fault-diagnosis test platform as the case study, 200 days of operating data were obtained from its 3-D model and the corresponding scaled test rig. Results show that the proposed framework triggers a yellow alarm on day 74, i.e. 25 days after the first soft-fault injection and 14 days after the second. After excluding the interpretable ∼10-day con- formation delay introduced by the soft-evidence stabilization module, the effective risk-response time is reduced to several days, and the framework escalates to a red alarm on day 78, providing a quantitative “when-to-intervene” decision basis for on-site predictive maintenance.

1. Introduction

The hoisting trolley system of an urban bridge erecting machine is the core actuator that performs the three motions of box-girder lifting, longitudinal travel and transverse travel. It carries the largest dynamic load of the whole machine and exhibits typical strong coupling among mechanical, hydraulic and electrical domains. This system operates frequently in densely populated urban areas. A failure

such as hoist over-speed or brake loss can easily lead to a girder-drop accident, causing severe casualties and property loss[1]. However, the hoisting trolley is a customized, single-piece, small-batch piece of equipment, and large-sample failure histories suitable for statistical inference are rarely available. Conventional reliability assessment methods built on a binary-state assumption are therefore not directly applicable, and small-sample reliability assessment based on multi-source information fusion has become one of the main development trends in the reliability field of large complex equipment[2].

Existing studies on reliability assessment for the hoisting trolley system follow two technical routes. The first is a data-driven fault-diagnosis route represented by ensemble learning algorithms such as XGBoost and LightGBM[3-5]. These methods extract non-linear features from low-level monitoring signals well, but are essentially cross-sectional classifiers and cannot describe the physical topological coupling among components or the accumulation of degradation risk over operating history. The second is a component-level system reliability reasoning route based on Bayesian networks (BN)[6-8]. These methods support both forward prediction and backward diagnosis along the topological dimension, but the prior inputs of their leaf nodes rely heavily on expert experience or offline statistics and are unable to handle high-dimensional continuous monitoring signals. In engineering practice, a long-standing gap exists between the data perception layer and the system reasoning layer of these two routes. In particular, for equipment such as the hoisting trolley dominated by progressive wear, fatigue and seal ageing, the conventional binary-state BN model cannot represent the soft-failure stage during the potential-to-functional (P-F) interval.

To address these issues, this paper proposes a real-time reliability assessment method based on a three state Noisy-OR Bayesian network with XGBoost evidence fusion. The main contributions are as follows: (1) the state space of bottom-layer BN nodes is extended from binary to a Normal/ Degraded/Failed three-state form, and a hybrid modelling strategy is adopted whereby mechanical/hydraulic components are described by three states and electrical/logical components remain binary; (2) a simplified three-state Noisy-OR risk propagation rule is constructed that preserves both the dominant “one-vote-veto” nature of hard failure and the cumulative effect of soft failure, while completing system-level inference with linear complexity aggregation formulas; (3) a soft-evidence injection mechanism is established that maps the XGBoost output to the BN leaf nodes, closing the loop between bottom-layer data perception and system level topological reasoning; (4) taking the hoisting-trolley fault-diagnosis test platform as the case study, 200 days of operating data are obtained from its 3-D model and the scaled test rig to quantitatively validate the alarm capability and response window of the framework in a predictive maintenance scenario.

2. Three-State Bayesian Network Model

2.1. Bayesian Network with Three-State Node Extension

A Bayesian network (BN) represents a set of random variables and their conditional dependencies using a directed acyclic graph. Each node carries a conditional probability table (CPT) conditioned on its parent set $Pa(X_i)$. The joint distribution of all N nodes can be factorized as[6]

$$P(X_1, X_2, \ldots, X_N) = \prod_{i=1}^{N} P(X_i \mid Pa(X_i)). \tag{1}$$

The bidirectional inference capability of the network naturally fits the coupled chain “sensor observation → component fault → system reliability”, because bottom-layer monitoring evidence can be propagated upward to estimate system risk, while system-level abnormality can also be traced backward to identify likely component causes.

The conventional binary BN restricts node states to "Normal/Failed" and cannot describe the progressive degradation and "operating with defects" P-F interval of core mechanical-hydraulic components in an urban bridge erecting machine (e.g. hoist wire ropes, slewing bearings and jacking cylinders). The state space of the bottom-layer nodes is therefore extended to three states in this paper: Normal ($X = 0$), Degraded ($X = 1$) and Failed ($X = 2$). Considering that electrical/logical units (relays, fuses, electromagnetic coils, etc.) typically exhibit Boolean abrupt-failure behaviour, forcing them into a "Degraded" middle state would have no physical meaning and would also cause an exponential explosion of the CPT size. Therefore, a hybrid modelling strategy is adopted: mechanical/hydraulic components are modelled with three states while electrical/logical components remain binary.

The prior probabilities of the bottom-layer three-state nodes are jointly calibrated from the catastrophic failure rate λ in a reliability database and the deterioration coefficient $\kappa = N_{\text{Degraded}} / N_{\text{Failed}}$ statistically estimated from the maintenance log:

$$P_F(t) = 1 - \exp(-\lambda t)\,, \tag{2}$$

$$P_D(t) = \kappa \cdot P_F(t)\,, \tag{3}$$

$$P_N(t) = 1 - P_F(t) - P_D(t)\,. \tag{4}$$

The corresponding 3 × 3 irreversible degradation prior-update matrix is[9]

$$\boldsymbol{A}(\Delta t) = \begin{pmatrix} e^{-(\lambda_d+\lambda_f)\Delta t} & \frac{\lambda_d}{\lambda_d+\lambda_f}\left(1 - e^{-(\lambda_d+\lambda_f)\Delta t}\right) & \frac{\lambda_f}{\lambda_d+\lambda_f}\left(1 - e^{-(\lambda_d+\lambda_f)\Delta t}\right) \\ 0 & e^{-\lambda_{df}\Delta t} & 1 - e^{-\lambda_{df}\Delta t} \\ 0 & 0 & 1 \end{pmatrix}, \tag{5}$$

where $\lambda_d = \kappa\lambda$ is the degradation transition rate, $\lambda_f = \lambda$ is the sudden-failure rate and $\lambda_{df} > \lambda_f$ is the soft-fault to hard-fault deterioration rate. The first row treats degradation and sudden failure as competing risks from the Normal state, so the row probabilities sum to one. In this paper $\boldsymbol{A}(\Delta t)$ acts as the prior ageing model of every three-state leaf node: starting from $[P_N(0), P_D(0), P_F(0)]^{\top}$ given by the catastrophic failure rate of the reliability database, the node prior for the current assessment day is updated by $p_i(t) = \boldsymbol{A}(\Delta t)^{\top} p_i(t-1)$ outside the BN topology and is then fused with the stabilized XGBoost evidence of that day through Eq. (12) before being aggregated upward by the Noisy-OR rule of Section 2.2. Thus, the Bayesian network topology remains static, while operating-time effects enter the assessment through age-dependent leaf priors and daily soft evidence.

2.2. Three-State Noisy-OR Risk Propagation

To describe the joint effect of multiple bottom-layer components on a parent node, a simplified three state Noisy-OR rule is adopted. Let states 0,1,2 denote Normal, Degraded and Failed, and let X_i be the children of parent Y. The rule consists of three points: if any child is in the Failed state, the parent enters the Failed state; if all children are Normal, the parent stays Normal; if no hard failure exists but at least one child is Degraded, the soft-failure influence propagates cumulatively according to an escape factor $q \in (0,1)$.

In online diagnosis, the stabilised evidence vector $\hat{\pi}_i$ is first obtained from the raw XGBoost output through the pipeline described in Section 4.4 and then fused with the age-dependent prior by Eq. (7). The resulting posterior belief vector is denoted $[P_{N,i}, P_{D,i}, P_{E,i}]$, and is used as the input to the NoisyOR aggregation. To meet the real-time computation requirement, the following aggregation formulas are adopted directly. Writing the parent-node state probabilities as a column vector, the three-state Noisy-OR aggregation can be expressed in the compact matrix form

$$\begin{bmatrix} P(Y=0) \\ P(Y=1) \\ P(Y=2) \end{bmatrix} = \begin{bmatrix} \prod_{i=1}^{n} P_{N,i} \\ \prod_{i=1}^{n}\left(P_{N,i}+qP_{D,i}\right) - \prod_{i=1}^{n} P_{N,i} \\ 1 - \prod_{i=1}^{n}\left(P_{N,i}+qP_{D,i}\right) \end{bmatrix}. \tag{6}$$

A smaller q makes a soft failure more likely to propagate to the parent node. The aggregation requires only two products pass over the n children of Y (one for $\prod_i P_{N,i}$ and one for $\prod_i\left(P_{N,i}+qP_{D,i}\right)$), giving an overall computational complexity of $O_{(n)}$ which is well suited to real-time reliability assessment.

3. XGBoost Soft Evidence Injection

3.1. XGBoost Algorithm

XGBoost (eXtreme Gradient Boosting) is adopted as the bottom-layer diagnostic model because it can handle nonlinear interactions among multi-source monitoring features and provide calibrated class probabilities through the multi:softprob objective [3]. In this paper, each three-state component node is assigned an independent XGBoost classifier. The classifier maps the sensor feature vector of the corresponding component to a three-class probability output associated with the Normal, Degraded and Failed states. Therefore, XGBoost is used here not as an independent endpoint classifier, but as the evidence-generation module that converts continuous monitoring data into probabilistic soft evidence for the Bayesian network.

3.2. From Classifier Output to Bayesian Soft Evidence

The three-class output of XGBoost is normalized by SoftMax into a raw probability vector $\pi_i^{\text{raw}} = \left[\pi_i^{(0)}, \pi_i^{(1)}, \pi_i^{(2)}\right]$, corresponding to the classifier scores that the i-th bottom-layer node is in the Normal, Degraded and Failed state respectively, with $\sum_k \pi_i^{(k)} = 1$. In the final online framework, this raw vector is not injected into the BN directly. It is first processed by the stabilisation pipeline in Section 4.4, and the resulting prototype vector $\hat{\pi}_i$ is used as the leaf-node soft evidence. Equivalently, the evidence likelihood is written as $L_i(X_i = k) = \hat{\pi}_i^{(k)}$. During inference, the BN multiplies it by the prior $P(X_i = k)$ and normalizes to obtain the posterior belief:

$$P'(X_i = k) = \frac{L_i(X_i=k)\cdot P(X_i=k)}{\sum_{j=0}^{2} L_i(X_i=j)\cdot P(X_i=j)}. \tag{7}$$

Unlike hard evidence that pins a node to a single deterministic state, the prototype soft evidence keeps a controlled amount of class uncertainty while filtering out the single-day jitter of the raw XGBoost output. This uncertainty is naturally fused into the system-level reasoning, leading to more robust assessment results.

3.3. Overall Framework

The overall method consists of four stages: “offline training → online diagnosis → soft-evidence stabilization → Bayesian feed back”. First, an XGBoost three-class model is trained offline using historical maintenance records and sensor data. During online operation, on each day a raw soft-evidence vector $\hat{\pi}_i^{\text{raw}}(t)$ is produced for every bottom-layer node. Before being injected into the BN, this raw vector is passed through the soft-evidence stabilization module (**EMA smoothing → dwell-confidence gate → prototype evidence injection**, detailed in Section 4.4) so as to filter out single-

day misclassifications and decision-boundary jitter. The stabilized vector, denoted $\hat{\pi}_i(t)$, is then injected into the corresponding BN leaf node, and the system-level three-state probabilities are inferred bottom-up using Eq. (6). The overall technical roadmap is shown in Fig. 1.

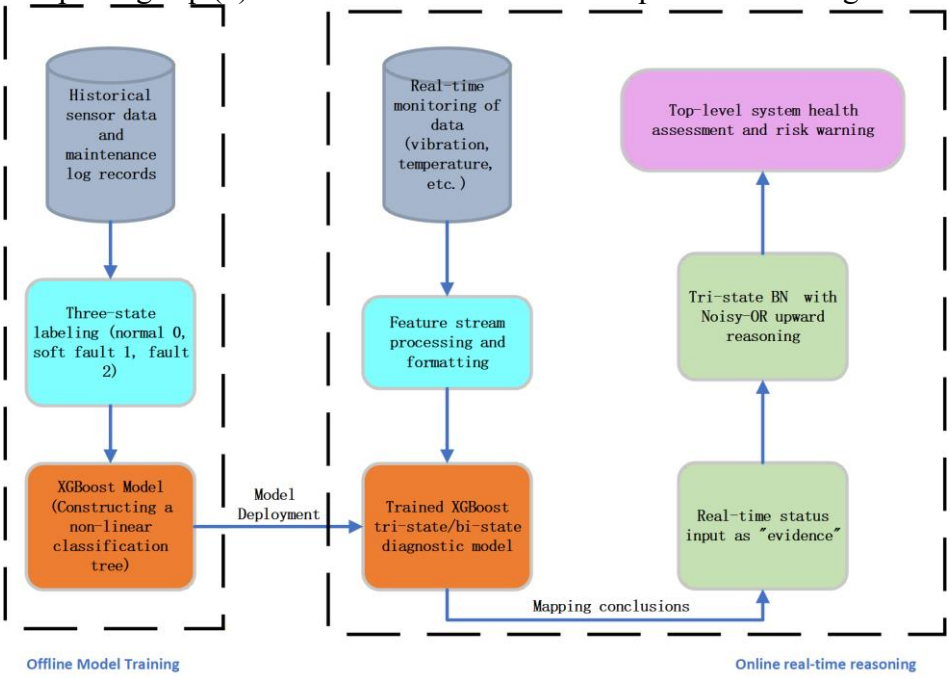

Figure 1 Real-time reliability assessment framework based on XGBoost soft evidence and the three-state Bayesian network

4. Case Study: 200 Days of Operating Data on the Hoisting Trolley Test Platform 4.1. Object Description and Sensor Features

The object of this study is the hoisting trolley system of an urban bridge erecting machine. Its Bayesian network contains 18 three-state mechanical/hydraulic nodes and 21 binary electrical/logical nodes. Every bottom-layer node is assigned a short identifier of the form EX#; in particular, EX16 denotes the hoist gearbox and EX2 denotes the wire rope, which are used as the two degradation injection targets in Section 4.5.

All training, validation and online assessment data come from a hoisting-trolley fault-diagnosis test platform composed of a 3-Dsimulation model (Fig.2)and a corresponding scaled physical test rig (Fig.3).

The two share the same sensing parameters and fault-injection interface, providing reproducible data for XGBoost training, soft-evidence injection and BN online inference.

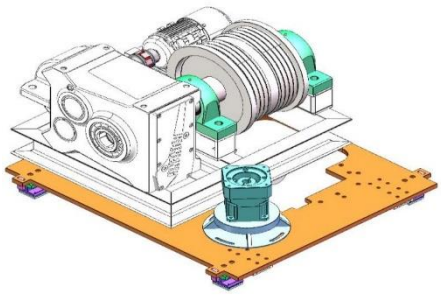

Figure 2 3-D simulation model of the hoisting trolley

Figure 3 Hoisting-trolley fault-diagnosis test platform of the bridge erecting machine

According to the topology described above, the Bayesian network of the hoisting trolley system is organized in a three-layer hierarchy: a top system node ("Hoisting Trolley System"), five intermediate subsystem nodes (Hoist, Brake, Long. Travel, Trans. Travel and Control & Power) and the 39 bottomlayerleafnodes (18mechanical/hydraulicthree-statenodesand21electrical/logicalbinarynodes). Causal edges between layers follow the Noisy-OR aggregation rule, and each leaf node receives its corresponding XGBoost soft evidence as input. The complete network structure is shown in Fig. 4.

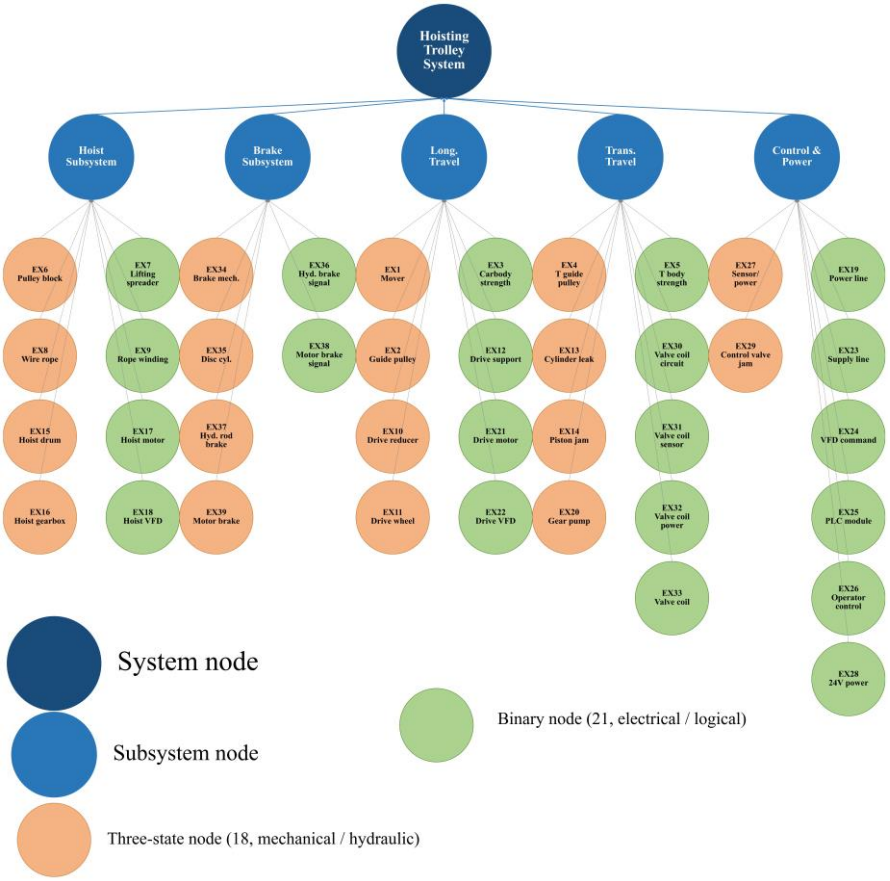

Figure 4 Bayesian network structure of the hoisting trolley system

4.2. Test Platform Dataset and XGBoost Training

It should be emphasized that the experiment reported in this section is one representative *case study* that demonstrates how the proposed framework can be deployed on a single bottom-layer node of the Bayesian network. The diagnostic target is intentionally chosen to be the **hoist gearbox (reducer)** rather than the variable-frequency hoist motor: as a typical mechanical-and-lubrication coupled component, the gearbox is dominated by progressive wear, gear-tooth pitting and lubrication degradation, which exhibits a clear "Normal → Degraded → Failed" P-F process and therefore best matches the three-state Noisy-OR modelling assumption of this paper. The motor itself is treated separately by another node of the network. The same procedure described below can be replicated on any other three-state node by replacing the corresponding sensor channels.

For the gearbox node, six categories of multi-source signals are continuously monitored on the test platform and used as XGBoost inputs: (i) **input-shaft torque / motor power** (power_w, kW), reflecting the load condition transmitted to the gearbox; (ii) **ambient wind speed** (wind_speed, m/s), used as a working-condition covariate of the trolley operating environment; (iii) **radial vibration amplitude on the gearbox housing** (rad_vib, mm/s, RMS in the 10–1000Hz band), the most

sensitive indicator of gear-pair and bearing wear; (iv) **ambient temperature** (env_temp, ◦C), used to compensate the absolute reading of the housing temperatures; (v) **drive-end (input-side) bearing temperature** (temp_drv, ◦C); and (vi) **non-drive-end (output-side) bearing temperature** (temp_nondrv, ◦C). Items (v) and (vi) jointly characterize the heat generation of the lubricated meshing pair and the bearings on both sides of the gearbox.

Three operating conditions are reproduced on the platform by physical fault injection on the gearbox: **Normal** (rated load, fresh lubricating oil, intact teeth); **Degraded** (slightly worn bearings, oil mixed with particulate contaminants – power and vibration amplitude rise by about 30% and the housing temperatures rise by about 10 ◦C compared with Normal); and **Failed** (tooth-spalled specimens installed and lubrication oil drained – power, vibration and temperature deteriorate simultaneously). The raw sensor streams are segmented by a fixed-length time window and normalized before being fed to XGBoost. The three classes overlap to some extent in the multi-dimensional feature space, reflecting the decision ambiguity present in real engineering situations.

A total of 7000 samples were collected with class ratio N_{Normal}:N_{Degraded}:N_{Failed} = 6000:600:400. The samples were split into training and test sets using stratified sampling at 7:3. The training set was then up-sampled by SMOTE to 4200 samples per class (12,600 balanced samples in total), and fed into an XGBoost three-class classifier (multi:softprob, K = 300, max_$depth$ = 4, η = 0.08, λ = 1, column subsampling 0.9).

4.3. Diagnosis Performance Verification

The diagnostic performance reported here is again specific to the *hoist gearbox node*: the trained XGBoost is evaluated on 2100 independent test samples produced by the gearbox fault-injection conditions described in Section 4.2. As summarized in Table 1, the AUC of all three classes is no less than 0.999, the macro-F1 reaches 0.979, and the recall of the Failed class is 0.95. The classifier therefore satisfies the credibility requirement for injecting gearbox-node soft evidence into the Bayesian network. Diagnosis models for the remaining bottom-layer nodes can be trained independently in the same way.

Table 1 Three-class performance of XGBoost on the test platform test set

Class	Precision	Recall	F1	AUC
Normal	1.0000	0.9989	0.9994	1.0000
Degraded	0.9570	0.9889	0.9727	0.9995
Failed	0.9828	0.9500	0.9661	0.9995
Macro avg	0.9799	0.9793	**0.9794**	0.9997

4.4. Operating Data Acquisition Scheme and Definition of the Warning Period

The data acquisition period is 200 days, with 12 h of work per day, a single step of $\triangle t$ =12h and a soft-channel escape factor q = 0.85. To reproduce the engineering scenario of "progressive degradation triggering an increase in system-level risk", two three-state nodes are randomly chosen on the platform and sequentially switched into the Degraded state: the first injection day $T_{\text{inj},1}$ is randomly drawn from [40,60] days, and after an interval $\triangle T \in [10,20]$ days the second node is injected. On each operating day t, according to the ground-truth state of each node, one sensor record is sampled from the corresponding operating-conditioning pool of the platform and fed into the trained XGBoost to produce the raw soft-evidence vector $\pi_i^{\text{raw}}(t)$ of that node for that day.

Soft-evidence stabilization. Direct injection of $\pi_i^{\text{raw}}(t)$ may amplify single-day sampling noise and decision-boundary misclassifications into repeated system-level threshold crossings. To suppress

this jitter, a three-stage stabilization pipeline "EMA smoothing → dwell-confidence gate → prototype evidence injection" is introduced:

1) Exponential moving average: $\tilde{\pi}_i(t) = \alpha\tilde{\pi}_i(t-1) + (1-\alpha)\pi_i^{\mathrm{raw}}(t)$, with $\alpha = 0.85$ to suppress single-day sampling fluctuations;
2) Dwell-confidence gate: let $s^*(t) = \mathrm{argmax}_k \tilde{\pi}_i^{(k)}(t)$ and let $\hat{s}_i(t-1)$ denote the currently confirmed state. The state switch $\hat{s}_i(t) \leftarrow s^*(t)$ is confirmed only if $s^*(t) \neq \hat{s}_i(t-1)$ and $\tilde{\pi}_i^{(s^*)}(t) - \tilde{\pi}_i^{(\hat{s}_i)}(t) \geq \delta$ (with $\delta = 0.15$) holds for $\tau = 10$ consecutive days; otherwise $\hat{s}_i(t-1)$ is retained;
3) Prototype evidence injection: the prototype vector $\bar{\pi}^{(\hat{s}_i)}$ is selected from below and sent into the BN as the soft evidence.

$$\bar{\pi}^{(0)} = [0.9995, 3*10^{-4}, 2*10^{-4}],$$
$$\bar{\pi}^{(1)} = [0.03, 0.94, 0.03],$$
$$\bar{\pi}^{(2)} = [0.02, 0.05, 0.93],$$

The stabilized vector is denoted $\hat{\pi}_i(t) \equiv \bar{\pi}_i^{(\hat{s}_i(t))}$. It is first fused with the age-dependent prior through Eq. (7), and the resulting posterior belief vector is then aggregated upward by Eq. (6). In this way, the BN responds to degradation confirmed by consecutive evidence rather than to isolated XGBoost fluctuations.

Different from the conventional whole-life endpoint defined by reliability dropping below a threshold for the first time, for the predictive-maintenance scenario of "soft-fault latency → system hard fault", a **system degradation warning period** is defined here as the interval from the first bottom-layer soft fault being *physically injected* (i.e. the ground-truth onset time $T_{inj,1}$) to the first crossing of the operational threshold by the system-level $P_F(t)$. This interval is the sum of two parts: the intrinsic confirmation delay of the stabilization module (about τ = 10 days) and the additional time needed for $P_F(t)$ to actually reach the threshold once the state switch has been confirmed. A two-tier threshold scheme is used: a yellow warning at $P_F \geq 0.30$ (a special inspection and shutdown preparation are recommended) and a red alarm at $P_F \geq 0.40$ (immediate shutdown maintenance is recommended):

$$T_{\mathrm{warn}}^{\mathrm{yellow}} = T(P_F \geq 0.30) - T_{inj,1}, \quad (8)$$
$$T_{\mathrm{warn}}^{\mathrm{red}} = T(P_F \geq 0.40) - T_{inj,1}, \quad (9)$$
$$T_{\mathrm{lead}} = T(P_F \geq 0.40) - T(P_F \geq 0.30). \quad (10)$$

4.5. Operating Results and Analysis

Following the above setup on the test platform, two faults are pre-injected to simulate a multiple progressive-degradation scenario: node 1 = EX16 (functional abnormality of the hoist gearbox) is injected on day $T_{inj,1}$ = 49, and node 2 = EX2 (wire-rope wear) is injected on day $T_{inj,2}$ = 60, 11 days after the first injection. The XGBoost model used for the EX2 wire-rope node is built with the same training procedure described in Section 4.2 (only the input channel set is adapted to the rope, with axial elongation and bending vibration replacing the gearbox-bearing temperatures), and reaches a comparable diagnostic level (macro-F1 ≈ 0.97); its detailed evaluation is omitted here for space. The system-level three-state probability and the stabilised soft-evidence trajectories of the two injected nodes are shown in Fig. 5; the numerical results at key time points are summarised in Table 2.

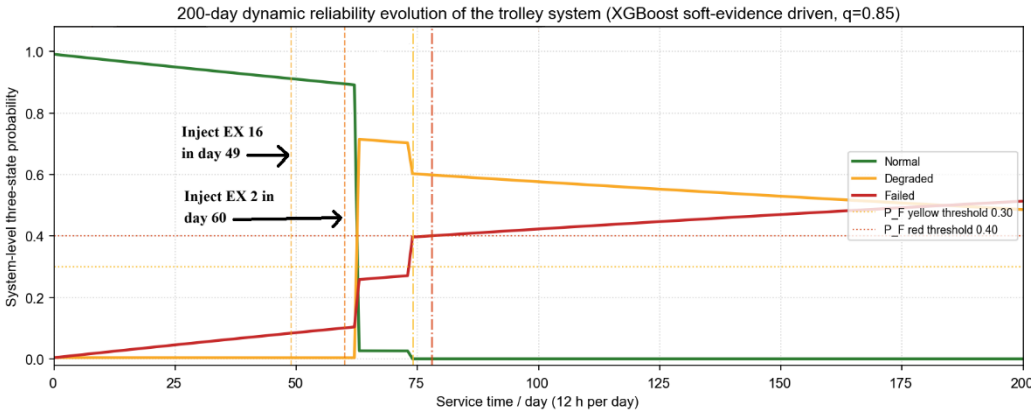

Figure 5 200-day operating results on the hoisting-trolley test platform

Table 2 System-level probabilities at key time points

Day	P_N	P_D	P_F	Stage description
0	0.9910	0.0046	0.0044	Initial commissioning state of the whole machine
20	0.9578	0.0044	0.0378	Cumulative drift of binary nodes
49	0.9116	0.0042	0.0842	EX16 soft-fault injection (not yet confirmed during dwell)
60	0.8947	0.0041	0.1012	EX2 injection (not yet confirmed during dwell)
74	0.0008	0.6465	0.3527	Both nodes confirmed in turn; yellow warning crossed
78	0.0008	0.5750	0.4242	Red-alarm threshold crossed
100	0.0008	0.5765	0.4228	Soft-fault steady-state holding
200	0.0006	0.4861	0.5132	Cumulative ageing of binary nodes drives a slow rise of P_F

The operating curves exhibit a three-stage characteristic. During days 0–48, all three-state nodes remain locked to the Normal prototype and the system-level P_F increases slowly below 0.10. During days49–73, EX16 and EX2 have physically entered the Degraded state, but their state switches are delayed by the dwell-confidence gate, which introduces an interpretable confirmation delay of about 10 days. After both nodes are confirmed, P_F rises rapidly, crossing the yellow-warning threshold on day 74 and the red-alarm threshold on day 78, and then remains in the high-risk band $P_F \approx 0.42 - 0.51$ without repeated threshold jitter.

The degradation warning periods computed from these curves are: $T_{\text{warn}}^{\text{yellow}} = 25\text{days}$,$T_{\text{warn}}^{\text{red}} = 29\text{days}$, with an available response window $T_{\text{lead}} = 4$ days. About 10 days of these come from the confirmation delay of the stabilization mechanism and represent a controllable price paid to eliminate "false-alarm jitter". Subtracting this confirmation delay, the system reaches the red-alarm threshold within several days after the confirmed coexistence of the two soft faults, without crossing the threshold back and forth, providing a quantitative and stable "when-to-intervene" basis for on-site predictive-maintenance decisions. In practice, the yellow-warning point can be interpreted as the beginning of an inspection-and-preparation window, during which the gearbox and wire-rope branches should be checked with priority and spare parts or shutdown resources can be arranged. The red-alarm point indicates that the accumulated soft faults have produced a system-level risk no longer suitable for continuous operation. Therefore, the 4-day interval between the two thresholds provides a practical buffer for changing from condition monitoring to maintenance execution. The exact warning period varies with the injected node positions, the injection interval and their coupling level in the subsystem topology. Nodes located in the same strongly coupled branch should therefore receive priority attention in maintenance planning.

5. Conclusion

To address the engineering issues that bottom-layer multi-source heterogeneous monitoring data and system-level probabilistic reasoning of an urban bridge erecting machine remain disconnected, and that conventional binary-state failure models cannot characterize progressive soft degradation, this paper proposes a real-time reliability assessment framework driven by "XGBoost soft evidence + a three-state Noisy-OR Bayesian network". The main conclusions are as follows:

(1) The state space of the bottom-layer BN nodes is extended from binary to a Normal/Degraded/Failed three-state form, and a hybrid modelling strategy whereby mechanical/hydraulic components use three states and electrical/logical components remain binary is adopted. This more accurately characterizes the "operating-with-defects" soft-failure stage during the P-F interval of the core mechanical-electrical-hydraulic components.

(2) A simplified three-state Noisy-OR risk-propagation rule is adopted; through Eq. (6), the posterior probabilities of bottom-layer nodes after soft-evidence fusion are aggregated layer by layer into the system-state probability. This preserves the rapid propagation of hard failure as well as the cumulative effect of soft failure and satisfies the computational requirements of real-time assessment.

(3) The case study shows that the SMOTE-balanced XGBoost submodule achieves macro-F1 = 0.979and AUC $\geq$ 0.999. On 200 days of operating data from the hoisting-trolley test platform, the framework triggers a yellow warning on day 74 and a red alarm on day 78, leaving a 4-day response window and satisfying the predictive-maintenance requirements of typical shutdown-overhaul cycles

Future work will further extend the source dimensions of soft evidence by introducing unstructured modalities such as images and acoustics, and incorporate maintenance action nodes into the Bayesian network to evaluate the lifetime reliability of the equipment under different maintenance policies.

Acknowledge

This study is supported by national key R&D plan of China (Grant No.2023YFC3806803)

References

[1] XU J, LI Z, WANG H, et al. Construction safety influencing factor analysis of bridge-erecting machines based on structural equation modeling[J]. Heliyon, 2024, 10(2).

[2] KHALEGHI B, KHAMIS A, KARRAY F O, et al. Multisensor data fusion: A review of the state-ofthe-art[J]. Information Fusion, 2013, 14(1):28-44.

[3] CHEN T, GUESTRIN C. XGBoost: A Scalable Tree Boosting System[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016:785794.

[4] ZHANG D, QIAN L, MAO B, et al. A Data-Driven Design for Fault Detection of Wind Turbines Using Random Forests and XGBoost[J]. IEEE Access, 2018, 6:21020-21031.

[5] TRIZOGLOU P, LIU X, LIN Z. Fault detection by an ensemble framework of Extreme Gradient Boosting (XGBoost) in the operation of offshore wind turbines[J]. Renewable Energy, 2021, 179: 945-962.

[6] WEBER P, MEDINA-OLIVA G, SIMON C, et al. Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas[J]. Engineering Applications of Artificial Intelligence, 2010, 25(4):671-682.

[7] BOUDALI H, DUGAN J B. A continuous-time Bayesian network reliability modeling, and analysis framework[J]. IEEE transactions on reliability, 2006, 55(1):86-97.

[8] CAI B, HUANG L, XIE M. Bayesian Networks in Fault Diagnosis[J]. IEEE Transactions on Industrial Informatics, 2017, 13(5):2227-2240.

[9] LANGSETH H, PORTINALE L. Bayesian networks in reliability[J]. Reliability Engineering & System Safety, 2007, 92(1):92-108.