A Graph Neural Network Approach to System-Level Health Index and Remaining Useful Life Estimation

Ark Ifeanyi[1]

[1] *University of Tennessee, Knoxville, TN, 37996, USA*
aifeanyi@vols.utk.edu

ABSTRACT

Current methods for predicting health index and remaining useful life (RUL) in complex systems struggle to account for performance dependencies between components, leading to inaccurate system-level estimates. This research proposes a novel approach utilizing graph neural networks (GNNs) to improve system-level health index and RUL estimation. GNNs excel at capturing complex interdependencies within a system, making them ideal for this task. The proposed methodology is designed for systems with synchronously sampled process data. To illustrate the application of the proposed approach, we will use the Condensate Extraction Subsystem (CES) of a nuclear power plant (NPP) as a case study. Sensor data like temperature, pressure, and flow rates will be used to train GNNs to predict the overall health and RUL of the CES over time. To evaluate the effectiveness of GNNs, a custom NPP simulator will be used to model the CES under various realistic fault modes across a variety of components. The GNN's performance will be verified and its robustness will be tested under diverse scenarios. This research aims to demonstrate the effectiveness and resilience of GNNs for system-level prognostics. By providing valuable insights for maintenance decision-making, this approach can enhance operational reliability and safety in complex engineering systems. The proposed framework has the potential to be applied across various industries, leading to advancements in predictive maintenance practices.

1. PROBLEM STATEMENT

In recent years, significant attention has been drawn towards prognostics and health management (PHM) techniques for predicting the remaining useful life (RUL) of complex systems. This paper aims to address the limitations of existing methods in integrating RUL information from individual components to derive accurate system-level RUL estimations. Current methodologies for RUL estimation often fall short when it comes to capturing the holistic health state of a system comprised of interconnected components. Some techniques employ direct mapping of system inputs to outputs to estimate system health or RUL, neglecting the nuanced degradation dynamics within individual components (Behera et al., 2021). Conversely, other methods focus on component-level prognostics, predicting RUL for each component and aggregating these predictions using tools like fault trees (Gomes et al., 2013). However, this second approach becomes computationally intensive for larger systems due to the necessity of establishing prognostic models for every individual component.

The complexity of system-level prognostics is further compounded by the intricate interdependencies between components, where the degradation of one component can influence and be influenced by others. These dependencies can lead to unique degradation patterns within the system, requiring a more sophisticated approach that accounts for these interactions (Kim et al., 2021). Moreover, the necessity of conducting system-level prognostics under various fault modes remains a critical challenge (Kim et al., 2021). Different fault modes can induce distinct degradation behaviors within the system, necessitating the development of predictive models capable of adapting to and accurately forecasting RUL across these diverse scenarios. Therefore, there is a clear imperative to develop novel methodologies that can effectively integrate RUL information from individual components, account for system-level interdependencies, and accommodate diverse fault modes to enhance the accuracy and applicability of system-level RUL estimation techniques.

This paper seeks to explore these challenges and propose a Graph Neural Network (GNN) based framework tailored towards system-level RUL estimation, leveraging the inherent relationships and dependencies between components to achieve more robust and accurate prognostic outcomes. By addressing these limitations, the research aims to contribute towards advancing the field of PHM and enabling more reliable maintenance strategies for complex engineering systems.

2. NOVEL CONTRIBUTION

To the best of our knowledge, the only application of GNN to system-level PHM in peer-reviewed literature focused on fault diagnostics (RuÍz-Tagle Palazuelos & Droguett, 2021). The researchers used a framework consisting of three key modules: (1) a system-level module that diagnoses and predicts the health of the entire system by leveraging learned embeddings from a graph convolutional network (GCN); (2) a component-level module that employs a deep graph convolutional network (DGCN) to diagnose individual component health states; and (3) a component interactions module based on a graph convolutional network autoencoder, which identifies and analyzes interactions among components during system degradation. The effectiveness of the framework was demonstrated through a case study involving a chlorine dioxide generation system (RuÍz-Tagle Palazuelos & Droguett, 2021).

Here, we attempt the first application of GNNs to system-level prognostics in an energy system with synchronous process data, considering all the necessary elements of system prognostics such as aggregation of components' degradation and interdependencies. Synchronous sampling ensures that all sensors collect data at the exact same moments, providing a perfectly aligned dataset for comprehensive analysis of simultaneous interactions. Furthermore, we propose a comprehensive way to verify the performance of the GNN models using three different methods. Finally, we suggest two separate ways to test the robustness of the proposed approach. The proposed method is demonstrated with a relatively simple system but is expected to be scalable to more complex systems and applicable across multiple industries.

3. RESEARCH PLAN

3.1. System-Level Solution

The proposed approach will be demonstrated on the subpart of a nuclear power plant (NPP) responsible for transporting the condensate from the condenser to the steam generator in a useful form and under the right conditions. This system referred to in the literature as the condensate and feed water system (CFWS), comprises the condensate extraction subsystem (CES) and the feed water subsystem (FWS) (Wang et al., 2016).

The CES functions to extract and pump the condensate from the condenser into the feed water subsystem. It typically includes a series of pumps designed to handle the condensate flow efficiently. Additionally, the CES often incorporates a low-pressure heater, which pre-heats the condensate before it enters the feed water subsystem (see Fig 1). This pre-heating helps optimize the overall energy efficiency of the system by utilizing waste heat from the condenser (Wang et al., 2016). The CES is the main subsystem of focus in this work given its crucial role in the power generation function of the NPP.

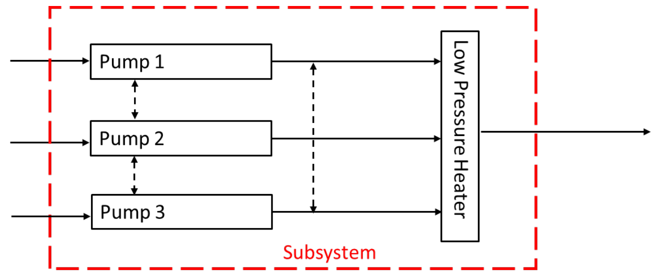

Figure 1. The Condensate Extraction Subsystem

The FWS receives pre-heated condensate from the CES and further processes it to prepare feed water for the steam generator. This subsystem includes pumps to pressurize the condensate and a high-pressure heater to increase its temperature to the required level. Together, the CES and FWS efficiently transform the condensate into high-quality feed water, optimizing the performance and efficiency of the nuclear power plant's steam generation cycle (Wang et al., 2016).

The main subsystem of focus is the CES but the entire plant is simulated using the Asherah NPP simulator (ANS) (Silva et al., 2020). ANS is a custom simulation tool designed specifically for performing cybersecurity assessments within nuclear power plants. It accurately simulates the operational behavior of a two-loop 2,772 MWt pressurized water reactor (PWR), including primary, secondary, and tertiary loops, as well as the control system. Developed using MATLAB/Simulink, ANS utilizes simple dynamic models for all its NPP components and systems (Silva et al., 2020). For this research, ANS will be further modified with relevant components for the system RUL estimation task. Specifically, pump degradation modules for two separate fault modes will be added to the CES subsystem and random noise will be added to the simulation to make sensor measurements more realistic.

Two specific fault modes to be simulated are "Loss of Net Positive Suction Head (NPSH)" and "Pump Blockage." The loss of Net Positive Suction Head (NPSH) occurs when a pump fails to maintain the necessary fluid pressure at its suction inlet, leading to the formation of cavitation—a phenomenon where vapor bubbles form and collapse within the pump due to low-pressure conditions. This cavitation may cause the pump to draw in air or gas along with the fluid, which can further impair operation. In severe cases, excessive cavitation can trigger automatic shut-off mechanisms or prevent the pump from starting due to insufficient fluid supply. (Lee et al., 2014). In the simulation, this behavior will be replicated by setting the pump's on/off command to 0 (or equivalent), indicating that the pump is not operating due to loss of NPSH. Pump blockage can lead to reduced flow rates or increased pressure within the pump system. This fault mode will be simulated by reducing the pump's speed (expressed as a percentage of maximum speed) to mimic decreased efficiency or capacity resulting from blockage. A blockage could be due to foreign objects in the pump or impeller fouling (Adams, 2017). The subsystem of interest shown in Fig 1 highlights that the input to the system is the input to the pumps and the system output is the output of the heater. The properties of the system to be monitored include the flow rates (kg/s), speed (%), differential pressure (Pa), enthalpy (J/kg), and temperature (K) of each CES component. Health parameters for each pump will be derived from these properties and their aggregated degradations will be monitored from the system output flows. The system RUL over time will be calculated from the system's health index and used as the target in one version of the experiment whereas in another version, the derived system health index will be the target variable.

The CES in an NPP presents a compelling case for utilizing GNNs, particularly temporal GNNs, for system-level RUL estimation. The strength of GNN lies in modeling the intricate relationships between components (Wu et al., 2020). Within the CES, the health of each pump directly impacts the others. A failing pump can reduce the flow rate, putting stress on the remaining pumps and potentially affecting pressure readings throughout the system. These interdependencies of the pumps are highlighted by the bi-directional dashed lines in Fig 1. GNNs excel at capturing these interdependencies (Wu et al., 2020), considering how changes in temperature, pressure, and flow rates from each pump influence the overall health of the CES and the Low-Pressure (LP) heater. In addition, temporal GNNs are adept at handling time-series data. Sensor readings from each pump, including temperature, pressure, and flow rates, could provide crucial insights into their health and degradation over time. Temporal GNNs can analyze how these metrics evolve (Cao et al., 2020), leading to more accurate RUL predictions. For example, the model could learn that a sudden drop in flow rate from one pump, coupled with a rise in temperature from another, might indicate a cascading fault mode. This capability to capture complex temporal dependencies between components, based on sensor data, makes temporal GNNs a potentially powerful tool for estimating the health state and RUL of the entire CES.

GNNs are employed to capture intricate system relationships by processing data in a structured manner (Wu et al., 2020). In this context, each component within the system is denoted as a node within the GNN framework. The interdependencies and interactions between these components are represented by edges that connect the corresponding nodes. Every node is associated with a feature vector containing specific information pertinent to that component, including sensor readings (such as temperature, pressure, and flow rate in the case of the CES), historical performance data, and other relevant attributes. Additionally, edges may possess features that indicate the nature or strength of the connections between connected nodes. Two principal input matrices are typically utilized by GNNs to process this data effectively (Wu et al., 2020):

Node Feature Matrix: This matrix is of dimension (number of nodes) x (feature vector size), where each row corresponds to the feature vector of a specific node in the system. For instance, in the CES context, each row in this matrix encapsulates sensor data or historical records associated with individual pumps or the LP heater.

Adjacency Matrix: The adjacency matrix captures the network's structure by representing node connections. This matrix is of dimension (number of nodes) x (number of nodes), where a value of 1 at a particular position (i, j) indicates a connection between node i and node j. The adjacency matrix can be binary (1 for connected, 0 for not connected) or weighted (reflecting connection strength). In the CES example, this matrix would illustrate the connections between pumps and the LP heater.

By processing these matrices, GNNs can discern how information propagates through the system, guided by the interconnections between components. The node features provide insight into individual component health and performance, while the adjacency matrix enables the GNN to grasp how these components influence one another (Wu et al., 2020). This holistic approach potentially empowers the GNN to deliver precise predictions regarding the overall health and RUL of the system.

3.2. Performance Verification

To evaluate the performance of the GNN model, three distinct approaches will be employed. Based on their recorded success in PHM for complex systems (Ifeanyi et al., 2024), initially, a deep neural network (DNN) model will be constructed and provided with all system inputs, including derived health indices of the components. This DNN will be trained to predict the target variable. Suppose the performance of this DNN matches that of the GNN, it suggests that the GNN might not be effectively utilizing the added information regarding component connections and interdependencies to make its predictions.

Secondly, a multi-DNN strategy will be implemented. One DNN will be designed to learn the degradation of an individual component, essentially mapping component inputs to their respective degradation or RUL. Simultaneously, another DNN will be developed specifically to understand how the individual degradations of components aggregate at the system level. This approach is challenging to scale due to the increasing number of DNNs required to train as the system's component count grows. If the GNN achieves comparable performance under this scenario, it will be favored due to its potential scalability.

Lastly, ablation studies will be conducted on the GNN model.

In these studies, specific connections within the GNN will be selectively removed, and the model's performance will be tested under these altered conditions. A significant decline in performance following connection removal would indicate that these connections significantly contribute to the GNN's predictive accuracy. This analysis will shed light on the importance of component connections in enhancing the GNN's effectiveness for prediction tasks.

3.3. Robustness Test

The effectiveness and robustness of the GNN will be rigorously tested across various fault modes and under conditions of data scarcity, with a particular focus on ensuring reliable performance under challenging scenarios. For the multi-fault-mode tests, the GNN will be evaluated under different fault configurations. This includes scenarios where a single pump experiences different faults sequentially, multiple pumps simultaneously experience the same fault type, and situations where multiple pumps concurrently exhibit different fault modes. By investigating these diverse fault scenarios, we aim to assess the GNN's capability to accurately diagnose and predict system degradation amidst complex operational challenges.

In terms of data scarcity, the GNN's performance will be assessed under decreasing sample sizes of input data. Initially, a large dataset will be used to train and test the model, ensuring optimal performance. Subsequently, the input sample size will be systematically reduced to examine the impact of limited data availability on the GNN's predictive capabilities. The study will identify the minimum sample size required for the GNN to maintain acceptable predictive performance, providing insights into the network's resilience against data constraints and its adaptability to real-world scenarios where comprehensive data may not always be readily available. These robustness tests are critical for validating the GNN's effectiveness and reliability as a predictive tool in practical applications. By systematically challenging the model with varying fault scenarios and limited data conditions, we aim to enhance the GNN's practical utility and ensure its applicability in dynamic and resource-constrained environments commonly encountered in industrial settings.

Expanding upon the investigation into data scarcity, future advancements in this research could involve augmenting available data using a sequence-to-sequence Variational Autoencoder (VAE). This approach would aim to enhance the quality and diversity of the dataset, particularly in scenarios where run-to-failure data is limited or unavailable. By leveraging the sequence-to-sequence VAE, which can generate synthetic data sequences representative of real operational conditions, we can enrich the dataset used for training the GNN. Subsequently, the augmented dataset's impact on the GNN's performance in predicting system-level degradation could be rigorously tested. This evaluation will provide valuable insights into the effectiveness of synthetic data augmentation techniques in improving the GNN's predictive capabilities under practical, real-world conditions characterized by data scarcity.

REFERENCES

Adams, M. L. (2017). *Power plant centrifugal pumps: problem analysis and troubleshooting*. CRC Press.

Behera, S., Misra, R., & Sillitti, A. (2021). Multiscale deep bidirectional gated recurrent neural networks based prognostic method for complex non-linear degradation systems. *Information Sciences*, *554*, 120–144.

Cao, D., Wang, Y., Duan, J., Zhang, C., Zhu, X., Huang, C., ... others (2020). Spectral temporal graph neural network for multivariate time-series forecasting. *Advances in neural information processing systems*, *33*, 17766–17778.

Gomes, J. P. P., Rodrigues, L. R., Galvão, R. K. H., & Yoneyama, T. (2013). System level rul estimation for multiple-component systems. In *Annual conference of the phm society* (Vol. 5).

Ifeanyi, A. O., Coble, J. B., & Saxena, A. (2024). A deep learning approach to within-bank fault detection and diagnostics of fine motion control rod drives. *International Journal of Prognostics and Health Management*, *15*(1).

Kim, S., Choi, J.-H., & Kim, N. H. (2021). Challenges and opportunities of system-level prognostics. *Sensors*, *21*(22), 7655.

Lee, S., Hassan, Y. A., Abdulsattar, S. S., & Vaghetto, R. (2014). Experimental study of head loss through an loca-generated fibrous debris bed deposited on a sump strainer for generic safety issue 191. *Progress in Nuclear Energy*, *74*, 166–175.

RuÍz-Tagle Palazuelos, A., & Droguett, E. L. (2021). System-level prognostics and health management: A graph convolutional network–based framework. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *235*(1), 120–135.

Silva, R., Shirvan, K., Piqueira, J. R. C., Marques, R. P., et al. (2020). Development of the asherah nuclear power plant simulator for cyber security assessment. In *Proceedings of the international conference on nuclear security, vienna, austria* (pp. 10–14).

Wang, H., Peng, M.-j., Wu, P., & Cheng, S.-y. (2016). Improved methods of online monitoring and prediction in condensate and feed water system of nuclear power plant. *Annals of Nuclear Energy*, *90*, 44–53.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Philip, S. Y. (2020). A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, *32*(1), 4–24.