Energy Consumption Prediction for Manufacturing in Industrial IoT Based on Heterogeneous GNN

Xuan Su*, Mingyan Li†‡, Qihao Li§, Chao Chen†, Songtao Guo†, Xu Wang*
* College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400044, P. R. China
† College of Computer Science, Chongqing University, Chongqing 400044, P. R. China
‡ Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, P. R. China
§ College of Communication Engineering, Jilin University, Changchun 130015, P. R. China
Emails: suxuan@stu.cqu.edu.cn, {limy2021, cschaochen, guosongtao}@cqu.edu.cn, qihaol@jlu.edu.cn, wx921@163.com

2024 IEEE/CIC International Conference on Communications in China (ICCC) | 979-8-3503-7841-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCC62479.2024.10681781

***Abstract*—Energy consumption prediction plays a crucial role in green manufacturing for industry, providing a theoretical foundation for energy-saving planning, workshop scheduling, and process optimization. However, owing to the processing complexity of mechanical products and the multiplicity of production information, previous methods can only achieve prediction for individual devices, which are quite restrictive to collaborative manufacturing planning. To deal with product-level energy management, this paper first proposes a heterogeneous manufacturing association graph (HetMG) which contains various manufacturing elements including product projects, decomposed orders and processing devices. Then, with HetMG, an energy consumption prediction method, HetEP, is introduced based on heterogeneous graph neural network (GNN), which incorporates Relational Graph Convolutional Networks (R-GCNs), and Long Short-Term Memory (LSTM) to mine the spatial-temporal relation of heterogeneous elements during the manufacturing process. Moreover, Graph Attention Network (GAT) is utilized in HetEP to adaptively learn the graph-structured data so that informative representations of each element can be obtained. The experimental results based on a manufacturing dataset from a generator factory demonstrate that the proposed HetEP yields superior prediction performance compared to the device-oriented prediction methods in terms of prediction accuracy.**

***Index Terms*—enery consumption prediction; green manufacturing; heterogeneous graph; deep learning network**

I. Introduction

With the growth of industrial activities in the context of industrial internet of things (IIoT), the total amount of carbon dioxide emitted every year is increasing [1]. Especially, the discrete manufacturing of large-scale mechanical products (e.g., turbine generators) with energy-intensive machinery is the key industry of energy consumption and carbon emissions. Due to concerns over climate change and the transition towards sustainable systems, a paradigm shift towards efficient and smart manufacturing systems is being promoted [2]. For example, the Union's carbon tariff policy mandates that companies specify energy consumption and carbon emissions per unit of product. Therefore, the metric of energy efficiency has recently gained significant attention.

For energy efficiency optimization, the calculation and prediction result of energy consumption can be used to aid process planning evaluation and workshop scheduling [3]. Hence, energy consumption prediction is the basis for fine energy management and the research of green manufacturing. The traditional energy consumption modeling method is mainly based on equipment mechanism analysis [4], [5]. However, the mechanism model cannot adapt to the real-time change of the production line state.

Nowadays, the advancement in sensor technology has enabled data acquisition in an energy-efficient and sustainable manner [6]. Meanwhile, due to the outstanding approximation capability of deep neural network (DNN), new progress has been made in data-driven methods for learning the complex energy consumption mechanisms of manufacturing systems [7]–[12]. Most existing research focused on device-oriented energy consumption prediction, where learning methods are used to analyze the potential mechanism of energy consumption in terms of machinery. More specifically, based on the prior knowledge of material flow and the historical energy consumption time series, the authors in [7] combined a graph convolutional network (GCN) and a gated recurrent unit (GRU) to predict the energy consumption of individual devices. By taking advantage of a large number of historical operation data, a rolling optimization strategy (ROS) was proposed in [8] to optimize the energy management strategy for the tram equipped with hybrid energy storage system. The work [9] proposed a computational DNN-based Multi-Gene Genetic Programming to model the energy consumption of a machine tool.

Moreover, a three-step framework using energy decomposition to predict machine loads was introduced in [10]. Similarly, an energy consumption model for a machine tool was created based on the cylindrical parts machining process in [11]. This model developed an method for each module and integrated them to form an overall model. Taking the multi-

This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFB3304500, and the National Natural Science Foundation of China under Grant 62203077, 52275476, and the Natural Science Foundation of Chongqing under Grant CSTB2022NSCQ-MSX1029, and the Talents Plan Project of Chongqing under Grant 0015094501001, and the Special Research Funding for Chongqing Postdoctoral Researchers under Grant 2021XM2015 and the Fundamental Research Funds for the Central Universities under Project NO.2023CDJXY-038, and the Open Project of Key Laboratory of System Control and Information Processing, Ministry of Education (Corresponding author: Mingyan Li).

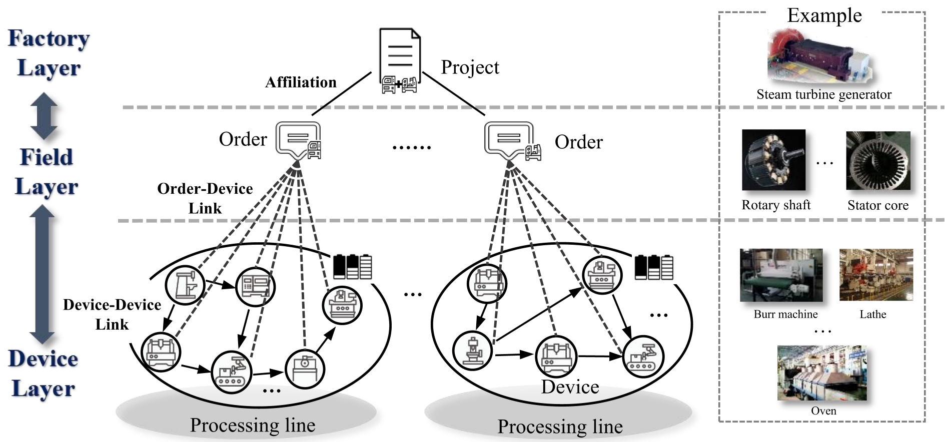

Fig. 1. Heterogeneous Manufacturing Association Graph (HetMG) with an example

axis coupling mechanism of a machine tool into account, an energy consumption model was developed under the multi-axis coupling condition, as described in [12]. Overall, much previous research investigates industrial energy consumption models based on learning techniques, where the parameters of the manufacturing process play a significant role. However, these efforts as well as the works on the traditional mechanism model normally predict energy consumption from only the device level and are quite restrictive to collaborative manufacturing planning from the entire enterprise perspective.

Challenge and Contribution: The development of industrial big data has made it possible to use comprehensive production information, offering a new opportunity to further improve energy efficiency from the upstream workshop and product levels. Nevertheless, the manufacturing system is formed by many interacting and interdependent components. As a result, a product's production energy consumption is closely tied to each workshop's devices and processing characteristics of product itself. Therefore, it is nontrivial to analyze the product-level energy consumption by simply using the relevant devices' consumption.

To deal with the product-level energy management, this paper presents a heterogeneous graph neural network (GNN) based energy consumption prediction method, namely HetEP, for mechanical product manufacturing. HetEP takes the relationship of manufacturing components as well as the production information of products as input and produces latent representations of all manufacturing elements including manufacturing devices, workshops, and products so that their energy consumption can be further obtained. The contributions of this paper can be summarized as:

- **Novel manufacturing association graph**: We first analyze the manufacturing elements of large-scale mechanical products, including product decomposition, workshop planning, and device scheduling in the manufacturing execution system (MES). Then, a heterogeneous graph topology is introduced to depict the explicit association of manufacturing elements, which are denoted as heterogeneous nodes with distinct attributes.
- **Spatial-temporal representation learning**: Based on the heterogeneous graph, the Relational Graph Convolutional Networks (R-GCNs) are used in HetEP to fuse various graph-structured production information to output latent representation of nodes. Moreover, the Long Short-Term Memory (LSTM) is utilized to mine the temporal underlying pattern of devices' historical energy consumption.
- **Attention-based energy consumption prediction**: To selectively aggregate production information from different kinds of manufacturing elements, two attention mechanisms are built in HetEP based on the heterogeneity of links between nodes. Finally, based on informative node representations, the energy consumption of each manufacturing element can be calculated and thus the targeted product-level prediction can be achieved.

The rest of this paper is organized as follows: Section II introduces the manufacturing association graph based on the manufacturing process, and Section III describes the details of HetEP. Section IV shows the experimental results, and Section V concludes this paper.

II. Heterogeneous Graph-based Hierarchy of Product Manufacturing Process

A. Discrete manufacturing Process of Mechanical Products

Large-scale mechanical product processing, e.g., electromechanical products, generally includes the production of raw materials, machining of parts, and assembly of components and complete machines. This complex manufacturing process consists of numerous large-sized parts with higher energy consumption compared to regular mechanical products. Hence, studying the energy consumed in each manufacturing process of large-scale mechanical products is of practical significance and can be used for energy-saving process optimization.

Typically, the production process for mechanical products in a factory is divided into three layers as follows.

- **Factory layer**: This layer is responsible for determining overall production tasks of products, e.g., using MOM

(the manufacturing operations management system), according to product requirements and processing status of the field layer.
- **Field layer**: The production tasks are then divided into several mechanical parts machining or assembling subtasks with their corresponding processing workshops.
- **Device layer**: In each field workshop, multiple devices or equipment, collaborate to complete the subtasks assigned based on production planning. Note that only the devices' energy-consumption power curves are accessible in realistic scenarios by deploying monitoring electricity meters based on IIoT systems.

More specifically, the factory layer uses MOM to realize production management of large-scale mechanical products. In MOM, the product processing task is called a project, the production manager will decompose the project (product) into parts, and then the processing task based on the parts will be sent to the field layer in the form of an order. In the device layer, the part manufacturing will be completed through local devices under multiple processing lines.

Example: Taking the manufacturing process of a turbine generator as an example, In the factory layer, the 60w turbine generator is decomposed into parts such as stator seat, stator core, coil, rotor shaft, end cap, bearing, etc. Subsequently, the processing of the stator seat, rotary shaft, and end cap are arranged in the metal and welding shops (field layer). In the metal shop, blanks are processed through the device layer of lathe, milling, and boring machines into the end caps.

B. Heterogeneous Manufacturing Association Graph

The energy consumption of a product is affected by multiple production processes at each layer. Due to the coupling manufacturing relationship among three layers, it is challenging to predict the energy consumption from the perspectives of factory and field layers with only devices' power curves. In this paper, we opt for a deep graph representation technique to formulate the energy-consumption knowledge model for discrete manufacturing systems. Since there are different kinds of components, the key step is to propose the novel manufacturing association graph based on a heterogeneous graph, referred to as HetMG.

To this end, the heterogeneous graph is first introduced, which is a special kind of information network that contains either multiple types of nodes or multiple types of links, formally defined in the following.

Definition II.1. ***Heterogeneous Graph.*** *A heterogeneous graph, denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, consists of a node set $\mathcal{V}$ and an edge set $\mathcal{E}$. A heterogeneous graph is also associated with a node type mapping function $\phi : \mathcal{V} \to \mathcal{A}$ and an edge type mapping function $\psi : \mathcal{E} \to \mathcal{R}$. $\mathcal{A}$ and $\mathcal{R}$ denote the sets of predefined node types and edge types, where $|\mathcal{A}| + |\mathcal{R}| > 2$.*

As shown in Fig.1, the objects of the three-layer factory architecture described above are abstracted as nodes $v \in \mathcal{V}$ in HetMG, and HetMG contains three kinds of nodes corresponding to the architecture, that is project, order, device, which are represented in HetMG as $\mathcal{A} = \{a_1, a_2, a_3\}$.

TABLE I
PRODUCTION INFORMATION

Layer	Node	Property
Factory	Project	Project_id
		Number of product
field	Order	Order_id
		Number of parts
		Material code
		Processing technology
Device	Device	Device_id
		Number of operation
		Processing time and part
		Energy comsumption

- The project node, i.e., $\phi(v) = a_1$, denotes a manufacturing planning of a set of the same products proposed by MOM. Each product manufacturing is divided into several parts machining.
- The order node, i.e., $\phi(v) = a_2$, is related to a typical part-oriented machining process and normally more than one part will be produced under an order. Note that the choice of operating parameter and material is the key factor of energy consumption for certain machine devices. Therefore, the material code and processing technology informations are collected.
- The device node, i.e., $\phi(v) = a_3$, refers to a manufacturing device in the device layer. The energy consumption of nodes calculated by the power curve can be available in the digital data acquisition and communication system along with the production information of processing time, parts, and the number of operations.

Each type of node in this factory-field-device framework, has a distinct property, including varying content of manufacturing information, listed as table I. This information is set as the feature $\mathbf{x}_v$ of node v, where the information for the device node contains time-varying energy consumption information.

Similarly, the edge $e \in \mathcal{E}$ between any two nodes captures the manufacturing relationship such as project to order, order to device, and device to device. Note that the link between project and order represents the process of decomposing products into components, and the product-level consumption can be obtained with the representation of nodes in the field layer. Therefore, we focus on the other edges, which are abstracted into different types defined as $\mathcal{R} = \{r_1, r_2\}$. r_1 represents the affiliation relation of devices and orders, r_2 is highly related to the process manufacturing and represents the precedence relation between devices in processing.

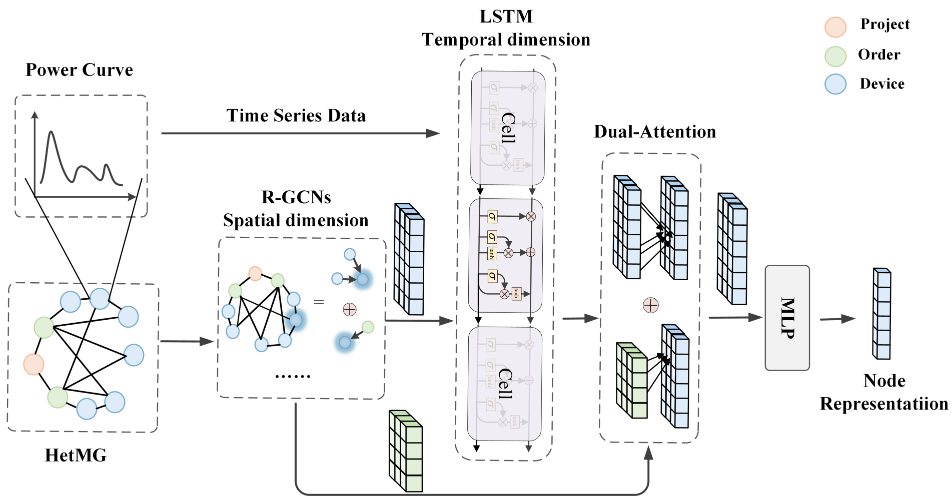

Fig. 2. The framework of HetFP based on heterogeneous manufacturing association graph

III. HetEP: Heterogeneous GNN-based Energy Consumption Prediction

In this section, HetEP, an energy consumption prediction method based heterogeneous spatial-temporal graph neural network is proposed for the discrete manufacturing of large-scale mechanical products. Fig. 2 shows the whole framework of HetEP, which contains three modules, i.e., R-GCNs, LSTM, and GAT (Graph Attention Network). From the spatial dimension, we integrate R-GCNs and GAT to realize the adaptive fusion of production information from heterogeneous neighboring nodes. From the temporal dimension, LSTM is utilized to capture dynamic evolution from time-power energy consumption sequences.

A. Spatial-Temporal Representation Learning

Recall that HetMG captures the relation of manufacturing process elements. The uncertainty of the manufacturing processing program has a significant impact on energy consumption, and thus the spatial interaction of nodes in HetMG is particularly important. Besides, compared with the homogeneous graph, the nodes of HetMG are heterogeneous with distinct features $\mathbf{x}_v, \forall v \in \mathcal{V}$.

Based on HetMG and these features, HetEP uses R-GCNs, a variety of GNN developed specifically to deal with the highly multi-relational data characteristic between heterogeneous nodes [13]. Based on R-GCNs, HetMG achieves the spatial integration of node information with different types as follows.

$$h_v^{k,a} = \sum_{u \in \mathcal{N}^a(v) \cup \{v\}} \frac{1}{\sqrt{\deg(v)} \cdot \sqrt{\deg(u)}} \cdot \left(W_a \cdot h_u^{k-1,a}\right) + b \tag{1}$$

where k, a denote the number of layers and the type of node, respectively. $\mathcal{N}^a(v)$ is the set of neighbor nodes of type a for node v and $\deg(\cdot)$ shows the degree of node. In each convolution shown as the kth layer, the updated node embedding $h_v^{k,a}$ can be got after integrating the information of neighbor nodes with type a, where $h_v^{0,a}$ is initialized to the node features $\mathbf{x}_v$. Note that, for each node type a, there is a separate weight matrix W_a for customized feature reconstruction and b denotes the bias.

Then, the node embeddings based on different types of neighbors are aggregated:

$$h_v^k = \text{ReLU}\left(\bigoplus_{a \in \mathcal{A}} \left\{h_v^{k,a}\right\}, h_v^{k-1}\right) \tag{2}$$

where h_v^k denotes the learned embedding of node v in the kth layer and the notation $\oplus$ denotes the aggregation scheme.

In addition to the spatial information in HetMG, the energy consumption history of devices in the manufacturing process is also important information, which is served in terms of time series. Therefore, after using R-GCNs for the spatial integration of each node, HetEP uses LSTM to capture the underlying varying patterns of the energy consumption sequence data.

For device node, i.e. $\phi(v) = a_3$, HetMG uses node embedding h_v^k in the last layer from (2) as input, together with the time series of energy consumption $x_v = \{x_v^0, x_v^1, \cdots, x_v^t\}$. Formally, the output of LSTM is defined as the node embedding, computed as,

$$\tilde{h}_v^t = \text{LSTM}(\tilde{h}_v^{t-1}, x_v^t; W_m) \tag{3}$$

where W_m is the weight matrix of the LSTM cell and is shared among all the device nodes.

B. Attention-based Energy Consumption Prediction

According to the principle of mechanical engineering, two factors determine the power curves of devices mainly. Firstly, the characteristic attributes of the product, such as its structural and functional features, determine processing techniques and parameters. These attributes indirectly influence the energy consumption of their processing devices in each manufacturing workshop in the form of order node's features. This impact can be reflected through the order-device links in HetMG. The second one is the processing sequence of the product,i.e., device-to-device links. This is because most devices in the workshop are substitutable and workers can flexiblely schedule these devices based on the current production situation. It will result in different devices being used to complete the processing tasks, which in turn leads to different energy consumption for the products. Therefore, to accurately learn the energy consumption models in both order and product levels, it is necessary to perform attentional training based on heterogeneous links with types in $\mathcal{R}$.

1) Order-Device Attention: HetEP uses the attention mechanism based on order-device links with type r_1, shown as,

$$h'_v = \text{GAT}\left(\tilde{h}_v, h_w; W_{r_1}\right) \tag{4}$$

where $\phi(v) = a_3$, $\phi(w) = a_2$, and $\psi(v, w) = r_1$. h_w is the order node's embedding obtained by R-GCNs in (2), while $\tilde{h}_v$ is the device node's embedding obtained by LSTM with time-varying data in (3). W_{r_1} denotes the weight matrix for the edge type r_1. Due to space limitation, the details of LSTM and GAT will not be discussed here, which can be seen in [14], [15], respectively.

2) Device-Device Attention: HetMG uses the latent information learned by spatial-temporal interaction to calculate the attention based on device-device links with type r_2, as

$$h''_v = \text{GAT}\left(\tilde{h}_v, \tilde{h}_w; W_{r_2}\right) \tag{5}$$

where $\phi(v) = a_3$, $\phi(w) = a_3$, and $\psi(v, w) = r_2$. W_{r_2} denotes the weight matrix for the edge type r_2.

Finally, HetEP aggregates the node respresentations obtained from (4) and (5), and pushes the updated node embedding into MLPs to get the final energy consumption.

$$\hat{y}_v = \text{MLP}\left\{\oplus\left(h'_v, h''_v\right)\right\} \tag{6}$$

Moreover, the energy consumption of the whole product can be obtained by the accumulation of the energy consumption of affiliation orders.

C. Training of HetEP

Based on the real values or labels of device energy consumption, the DNN parameters of HetEP can be trained by the root-mean-square error (RMSE) calculated as,

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2} \tag{7}$$

where $\hat{y}$ and y denote the predicted and real energy consumption values, respectively. N denotes the number of devices involved.

IV. EXPERIMENTS

In this section, we take the generator rotor as an example and use various information during its manufacturing process to predict energy consumption.

A. Datasets

The main production information about the generator rotor is shown in Tab II. This information is collected in MOM and MES of the factory. The energy consumption dataset of devices is collected and obtained by the digital data acquisition and communication system.

TABLE II
ROTOR PROCESSING INFORMATION

Workshop	**Component**	**Manufacturing Process**	**Device**
Metal Workshop	shaft	Turning	01004
		Milling	01007,01008
		Drilling	01010,01012
Welding Workshop	coil	sawing	05008
	collector ring	welding	05001
Coil Workshop	coil	winding	03001,03002

The data set contains the energy consumption values of each device during the execution of rotor manufacturing tasks, with the time range from 02/12/2023 00: 00 to 09/12/2023 23: 59. From this data set, 13 560 are used for model training, 10 520 for testing, while 10 520 are validation observations.

HetEP converts the processing information shown in Tab II into features of nodes. For textual information, it is converted by one-hot encoding. Digital information is determined according to the regulations of the manufacturing industry or the coding rules developed within the factory, so it already has specific semantics and need not be converted.

B. Baseline Models

The following baseline models are used to compare the performance with HetEP:

(1) GRU: As an improved RNN, the GRU in [16] has a strong capacity for capturing temporal relationships from time series. Thus, the GRU is chosen to construct the temporal relationship learning network.

(2) ConvLSTM: The ConvLSTM in [17] has convolutional strutures in both the innput-to-state and state-to-state transitions, so it can captures spatiotemporal correlations better.

(3) GC-LSTM: Proposed in [18], in this model, GCN is capable of node structure learning of network snapshot for each time slide, while LSTM is responsible for temporal feature learning for network snapshot.

C. Experimental Results

As can be seen from Tab.III, the proposed HetEP method outperforms other models in all metrics by capturing both spatial and temporal features. The GC-LSTM performs satisfactorily, but the HAGAT method is still better.

D. Ablation Study

To validate the effectiveness of key components of HetEP, we conduct ablation studies, primarily investigating the impacts of the R-GCNs module as well as the GAT module on

TABLE III
THE PREDICTION PERFORMANCES OF METHODS

Method	RMSE	MAE	sMAPE
HetEP	**3.255**	**2.940**	**18.7%**
GRU	8.965	10.354	25.6%
ConvLSTM	9.579	8.732	23.4%
GC-LSTM	5.325	3.251	22.7%

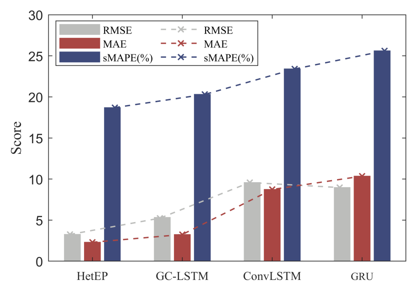

Fig. 3. Exprimental Results

the final results. We remove the R-GCNs module and GAT module separately and use GC-LSTM as the baseline.

TABLE IV
ABLATION STUDY

Module		Evaluation Metrics		
R-GCNS	GAT	RMSE	MAE	sMAPE
		5.325	3.251	22.7%
✓		4.851	3.012	20.4%
	✓	4.846	3.217	21.9%
✓	✓	**3.255**	**2.940**	**18.7%**

The results are shown in Tab.IV. HetEP with R-GCNs improves the results as it makes good use of the heterogeneity of nodes to realize the processing of heterogeneous information in the manufacturing system. The impact of GAT can be observed in Tab.IV, which confirms that employing an attention mechanism is advantageous since it allows for the selective emphasis on the concealed information present between various node types.

V. CONCLUSION

In this paper, we established an energy consumption prediction method namely HetEP for green manufacturing, which can predict energy consumption in order and product levels based on heterogeneous graph neural networks. The heterogeneous topology graph HetMG is firstly proposed to capture the product processing flow. Based on HetMG, HetEP leverages R-GCNs to fuse spatial heterogeneous node information, and then uses LSTM to mine the potential patterns among the energy consumption data of devices. Finally, two kinds of attention mechanisms for the heterogeneity of links are proposed in HetEP to differentiate the importance of nodes under different links. In the future, we aim to integrate node representation learning and processing scheduling to further realize green manufacturing.

REFERENCES

[1] Yanjiu Zhong, Jingcheng Wang, Jiahui Xu, Shunyu Wu, Jun Rao, and Kangbo Dang. Prediction of energy consumption in horizontal roughing process of hot rolling strip based on TDADE algorithm. *IEEE Transactions on Automation Science and Engineering*, 21(1):555–568, 2024.

[2] Aicha Dridi, Chérifa Boucetta, Hassine Moungla, and Hossam Afifi. Deep recurrent learning versus q-learning for energy management systems in next generation network. In *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021.

[3] Milad Mohammadi, Song Han, Ehsan Atoofian, Amirali Baniasadi, Tor M. Aamodt, and William J. Dally. Energy efficient on-demand dynamic branch prediction models. *IEEE Transactions on Computers*, 69(3):453–465, 2020.

[4] Markus Brillinger, Marcel Wuwer, Benjamin Smajic, Muaaz Abdul Hadi, Stefan Trabesinger, Bernhard Oberegger, and Markus Jäger. Novel method to predict the energy consumption of machined parts in the design phase to attain sustainability goals. *Journal of Manufacturing Processes*, 101:1046–1054, 2023.

[5] Junhua Zhao, Li Li, Lingling Li, Yunfeng Zhang, Jiang Lin, Wei Cai, and John W Sutherland. A multi-dimension coupling model for energy-efficiency of a machining process. *Energy*, 274:127244, 2023.

[6] Muhammad Moid Sandhu, Sara Khalifa, Raja Jurdak, and Marius Portmann. Task scheduling for energy-harvesting-based iot: A survey and critical analysis. *IEEE Internet of Things Journal*, PP(99), 2021.

[7] Jianhua Guo, Mingdong Han, Guozhi Zhan, and Shaopeng Liu. A spatio-temporal deep learning network for the short-term energy consumption prediction of multiple nodes in manufacturing systems. *Processes*, 10(3):476, 2022.

[8] Yu Wang, Zhongping Yang, Fei Lin, Xingkun An, and Xiaochun Fang. A hybrid energy management strategy based on line prediction and condition analysis for the hybrid energy storage system of tram. *IEEE Transactions on Industry Applications*, PP(99):1–1, 2020.

[9] Shailendra Pawanr, Girish Kant Garg, and Srikanta Routroy. Prediction of energy consumption of machine tools using multi-gene genetic programming. *Materials Today: Proceedings*, 58:135–139, 2022.

[10] Daniel Tan, Manu Suvarna, Yee Shee Tan, Jie Li, and Xiaonan Wang. A three-step machine learning framework for energy profiling, activity state prediction and production estimation in smart process manufacturing. *Applied Energy*, 291:116808, 2021.

[11] Shailendra Pawanr, Girish Kant Garg, and Srikanta Routroy. A novel approach to model the energy consumption of machine tools for machining cylindrical parts. *Journal of Manufacturing Processes*, 84:28–42, 2022.

[12] Jinwen Zhang, Congbo Li, Yongsheng Li, Ningbo Wang, and Wei Li. Energy consumption modeling and optimization of a hobbing machine tool considering multi-axis coupling. *IEEE Transactions on Automation Science and Engineering*, 2023.

[13] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *The semantic web: 15th international conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, proceedings 15*, pages 593–607. Springer, 2018.

[14] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[15] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

[16] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.

[17] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28, 2015.

[18] Jinyin Chen, Xueke Wang, and Xuanheng Xu. Gc-lstm: Graph convolution embedded lstm for dynamic network link prediction. *Applied Intelligence*, pages 1–16, 2022.