MDPI

Article

Design of a Transformer-GRU-Based Satellite Power System Status Detection Algorithm

Guoqi Xie 🆔 **, Xinhao Yang *, Jiayu Zhao and Zhou Huang**

School of Mechanical and Electrical Engineering, Soochow University, Suzhou 215006, China; 20234229040@stu.suda.edu.cn (G.X.); 20215229014@stu.suda.edu.cn (J.Z.); 20235229129@stu.suda.edu.cn (Z.H.)
* Correspondence: yangxinhao@suda.edu.cn

Abstract

The health state of satellite power systems plays a critical role in ensuring the normal operation of satellite platforms. This paper proposes an improved Transformer-GRU-based algorithm for satellite power status detection, which characterizes the operational condition of power systems by utilizing voltage and temperature data from battery packs. The proposed method enhances the original Transformer architecture through an integrated attention network mechanism that dynamically adjusts attention weights to strengthen feature spatial correlations. A gated recurrent unit (GRU) network with cyclic structures is innovatively adopted to replace the conventional Transformer decoder, enabling efficient computation while maintaining temporal dependencies. Experimental results on satellite power system status detection demonstrate that the modified Transformer-GRU model achieves superior detection performance compared to baseline approaches. This research provides an effective solution for enhancing the reliability of satellite power management systems and opens new research directions for future advancements in space power system monitoring technologies.

Keywords: power system; transformer; gated recurrent unit; graph attention network; state detection

1. Introduction

The space industry has exerted multifaceted impacts on human society, particularly through artificial satellites that generate substantial economic value. The adverse electromagnetic environment in outer space poses significant threats to satellite safety. Among the satellite platform's subsystems, the power system experiences the highest failure rate. Satellite power system failures account for 28% and 26% of the total spacecraft anomalies domestically and internationally, respectively. Among 249 domestic power system anomalies reported between 2007 and 2017, 97.2% were minor faults recoverable in orbit. The most prevalent fault types in energy storage battery systems include performance degradation anomalies, short-circuit faults, open-circuit faults in individual cells, and abrupt overcharging/deep discharge events caused by power control equipment failures. Failure statistics further reveal that energy storage battery systems account for approximately 40% of all power system failures, followed by power control equipment-related faults at 21% [1]. Therefore, the rapid and accurate detection of the health state of energy storage battery packs within the satellite power system is of critical importance for ensuring the long-term, stable operation of satellite platforms.

Lithium-ion batteries are widely employed in the energy storage battery packs of satellite power systems due to their advantages such as high energy density and long service life [2]. State detection methods for lithium-ion batteries are primarily categorized into model-based approaches [3–6] and data-driven approaches [7,8]. While model-based methods (e.g., state observers combined with electrochemical models [9] or equivalent time-varying internal resistance models [10]) offer strong interpretability, they require high model accuracy and real-time performance, making them less adaptable to complex operational scenarios or situations without precise modeling [10]. By contrast, data-driven approaches eliminate the need for exact modeling and demonstrate superior adaptability and potential by autonomously learning data patterns. In recent years, deep learning techniques have been extensively applied in this field, such as support vector machines for voltage anomaly diagnosis [11] and deep neural networks for lifespan prediction [12]. Notably, Transformer models have exhibited exceptional performance in time-series prediction tasks due to their capability to capture long-range sequence dependencies. For instance, Huang Yufeng et al. [13] achieved promising results in predicting lithium battery health states by integrating a Kalman filter-optimized Transformer network.

However, existing Transformer-based methods applied to satellite power system state detection still face significant challenges:

(1) Inadequate spatial correlation modeling: Satellite energy storage battery packs consist of multiple interconnected battery cells arranged in series, where inter-cell states are mutually influenced by physical mechanisms such as thermal coupling and fault propagation [1,14]. While the self-attention mechanism in standard Transformers can capture internal sequence dependencies, it fails to explicitly model the inherent spatial correlations induced by the physical connectivity (graph structure) among battery cells. Current approaches (e.g., in [13]) typically treat the battery pack as a monolithic unit or consider individual cells independently, neglecting critical topological information.

(2) Computational efficiency and real-time constraints: The Transformer decoder architecture is relatively complex and computationally expensive [15]. Satellite in-orbit applications impose stringent real-time requirements on algorithms, making traditional Transformer decoders potentially inadequate for rapid state detection tasks.

(3) The optimization of long-term temporal pattern modeling: Satellite telemetry data (e.g., temperature) often exhibit periodic variations (e.g., Earth eclipse cycles [16]) and demonstrate time-dependent fault evolution characteristics. While the Transformer excels at modeling long-range dependencies, its positional encoding schemes and attention mechanisms still require refinement to more effectively capture these strong periodic patterns and specific temporal dynamics in satellite data [13,16].

(4) Satellite in-orbit data are constrained by communication bandwidth limitations, resulting in a scarcity of complete charge/discharge cycle datasets. Fault samples, particularly early progressive faults, account for less than 0.1% of the data. This data sparsity hinders models from capturing subtle patterns of fault evolution, posing significant challenges for anomaly detection and predictive maintenance in satellite power systems.

To address the aforementioned challenges, this paper proposes an enhanced Transformer-GRU model (TF-GRU) [17] integrating a Graph Attention Network (GAT) and a GRU for satellite power system status detection. Specifically, a GAT layer [18] is embedded within the Transformer encoder to explicitly model spatial correlations among battery cells (treated as graph nodes). The GAT dynamically learns attention weights between nodes, enabling the adaptive capture of fault propagation paths and influence ranges within the battery pack (e.g., how cell faults affect neighboring batteries [14]). This significantly enhances the model's capability to extract spatially correlated features, overcoming the lim-

itations of standard Transformers in this regard. This approach draws inspiration from the successful application of graph neural networks in handling non-Euclidean relationships (e.g., networks and knowledge graphs) [19], adapting it to battery packs with well-defined topological structures. The GRU network [20], known for its simple yet efficient cyclic architecture, replaces the conventional Transformer decoder. GRUs effectively capture short-term dependencies and dynamic patterns in temporal sequences, making them particularly suitable for predicting satellite parameter trends (e.g., charge/discharge voltage curves and periodic temperature variations). This substitution significantly reduces the model's computational complexity, meeting the stringent real-time requirements of on-orbit satellite status detection while preserving temporal dynamic modeling capabilities. The improved TF-GRU model achieves effective spatiotemporal feature fusion by leveraging the GAT-enhanced encoder for spatial correlation extraction and the GRU decoder for efficient temporal dynamics processing. This design aims to comprehensively uncover hidden patterns in satellite power system data, particularly abnormal patterns associated with fault events.

This study selects readily available voltage and temperature data from battery packs as model inputs [1,16,18], and a series of experiments—including trend prediction and fault detection—are conducted to validate the effectiveness of the proposed model. The experimental results demonstrate that the improved Transformer-GRU model not only accurately predicts the parameter variation trends of satellite power systems but also promptly and effectively detects faults in energy storage battery cells (e.g., short circuits and open circuits) and power control equipment (e.g., failure of charge regulators). The model outperforms conventional Transformer and standard TF-GRU models in both detection accuracy and efficiency. This research provides a more effective technical solution for enhancing the reliability of satellite power management systems.

2. Materials and Methods

The data employed in this study were collected from a geostationary orbit (GEO) communication satellite, comprising telemetry measurements of voltage and temperature from the onboard energy storage battery system during its operational mission. The satellite's 15-year design lifetime, coupled with its 24 h orbital period, subjects the battery to a stable illumination–shadow cycle. This operational profile contrasts sharply with the high-frequency charge/discharge patterns typical of low Earth orbit (LEO) satellites. The battery system utilizes a customized non-commercial design, featuring a series-connected topology and dedicated thermal management architecture proprietary to aerospace engineering. Notably, detailed technical specifications remain confidential due to supplier intellectual property constraints. The data were sampled at a frequency of 10 Hz and spanned multiple umbra periods (when the satellite entered the Earth's shadow region). During the data preprocessing phase, abnormal data points exceeding the voltage thresholds (64.3–72.5 V) and temperature thresholds (17–40 °C) were first excluded using outlier detection algorithms (e.g., Z-score or IQR methods). Missing data caused by communication interruptions or sensor malfunctions were imputed using mean imputation techniques. Subsequently, the voltage and temperature data were normalized to the [0, 1] interval via Min–Max normalization. Finally, the preprocessed data were chronologically partitioned into training and testing sets. The training set comprised 71,722 samples, spanning multiple operational cycles (including normal conditions and minor anomalies), and it was 16 times larger than the test set (4490 samples). The test set specifically captured complete charge/discharge cycles of the satellite battery, encompassing dominant operational modes such as discharging during eclipse periods and charging during sunlit intervals. This dataset configuration ensured comprehensive model validation across representative in-orbit scenarios.

In terms of model design, this study proposes an improved Transformer-GRU model (TF-GRU) that integrates a GAT with a GRU. The Transformer model consists of an encoder and decoder, with its core mechanism being the self-attention mechanism, which enables the capture of global dependencies in input sequences. The encoder architecture incorporates multi-head self-attention modules and feedforward neural networks, while residual connections and layer normalization are employed to mitigate gradient vanishing issues. Positional encodings are generated using sine and cosine functions to encode fixed positional information. The GRU network regulates information flow through an update gate and a reset gate, computing the final state based on the candidate state. To enhance the model's ability to capture spatial correlations, a GAT layer is embedded within the Transformer encoder to dynamically adjust attention weights among battery cells. In a battery pack comprising N individual cells, each cell is treated as a graph node. The input to the model is a set of attention feature vectors $h = \{\vec{h}_1, \vec{h}_2, \cdots, \vec{h}_N\}$, $\vec{h}_i \in R^F$ for all nodes, and the output corresponds to the updated attention feature vectors $h' = \{\vec{h}'_1, \vec{h}'_2, \ldots, \vec{h}'_N\}, \vec{h}'_i \in R^{F'}$. The correlation between an arbitrary node i and its neighbor j is quantified by an attention coefficient a_{ij}, defined as follows:

$$a_{ij} = \frac{exp\left(LeakyReLU\left(\vec{a}^T\left[w\vec{h}_i \parallel w\vec{h}_j\right]\right)\right)}{\sum_{k \in N_i} exp\left(LeakyReLU\left(\vec{a}^T\left[w\vec{h}_i \parallel w\vec{h}_k\right]\right)\right)} \tag{1}$$

where *LeakyReLU* denotes the nonlinear activation function, *a* represents the self-attention weight vector, w is the learnable weight matrix, $\parallel$ denotes the concatenation operation, and N_i denotes the set of neighboring nodes for node i.

The updated output of node ii is obtained by performing a weighted summation over the feature vectors of all its neighboring nodes i. Additionally, the GAT introduces a multi-head mechanism, where K independent attention heads are computed. Their outputs are aggregated through concatenation or averaging operations to form a new feature representation, thereby enhancing the model's capacity to extract critical features and improving training stability.

This study proposes a hybrid architecture that integrates a GAT into the Transformer encoder to explicitly model spatial correlations among battery cells (e.g., thermal coupling and fault propagation pathways), overcoming the limitations of pure Transformer architectures in capturing physical topological relationships. The conventional Transformer decoder is replaced with a GRU enabling efficient temporal dynamics modeling through its gating mechanisms and addressing the inefficiency of pure GRUs in capturing long-term dependencies. The integration of a GAT and GRU achieves synergistic spatial feature extraction and temporal sequence analysis, preserving the Transformer's capability for long-range dependency modeling while reducing computational complexity. This framework optimizes dynamic weight allocation for spatial features and memory updates for temporal patterns in satellite battery data. By incorporating physical graph structure constraints and temporal gating mechanisms, the model maintains real-time performance while avoiding the global attention redundancy of pure Transformers and the temporal information loss inherent to pure GRUs. The improved TF-GRU model architecture is illustrated in Figure 1. The model architecture is defined by an input dimension of 3, corresponding to voltage, temperature, and temporal features. To accommodate the long temporal characteristics of satellite data, positional encoding is implemented with a maximum sequence length of 256. The attention mechanism employs 4 heads combined with causal attention masking to model temporal dependencies. A GRU layer comprising 128 hidden units is followed by a ReLU activation layer and a dropout layer with a 1% dropout rate to mitigate over-

fitting. The output layer consists of a three-neuron fully connected layer, generating a three-dimensional output vector representing voltage, temperature, and battery pack state.
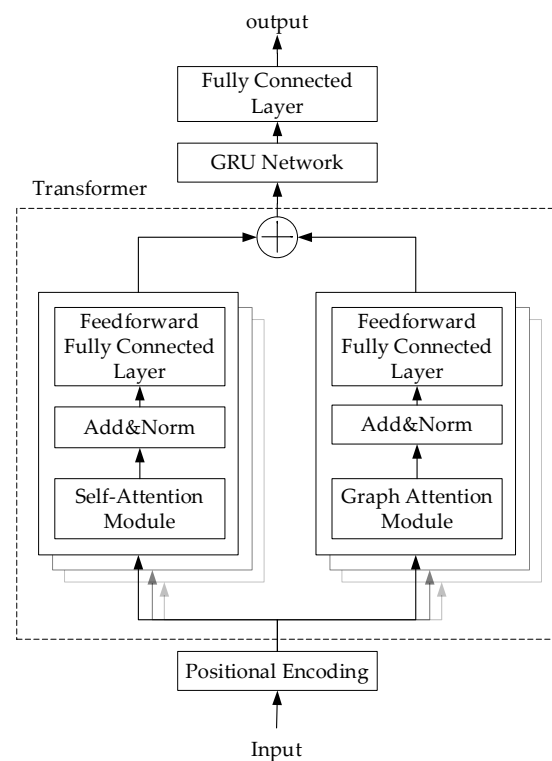
Figure 1. Schematic diagram of the improved Transformer-GRU model architecture.

To validate the effectiveness of the improved model, comparative experiments were conducted against the conventional Transformer model and the standard TF-GRU model. The performance of the models was quantified using the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE).
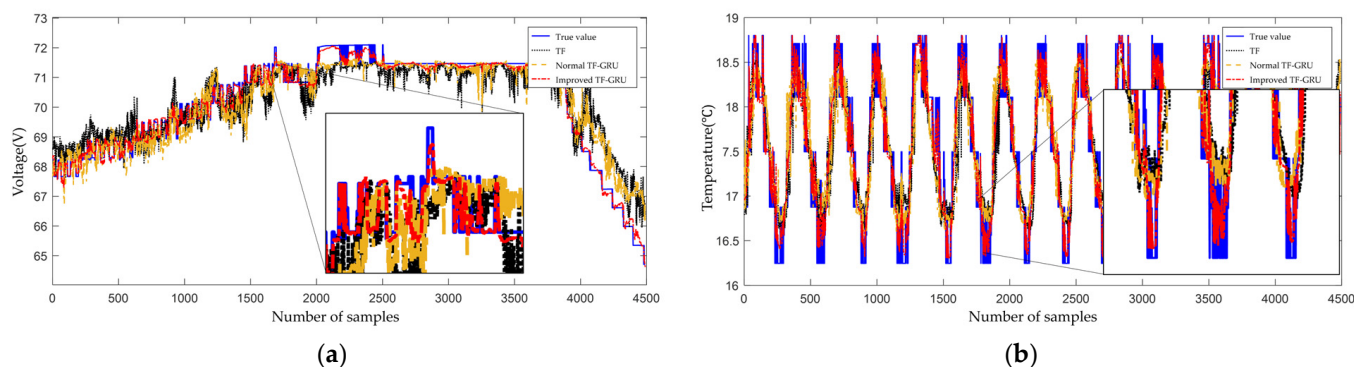
3. Results

3.1. Satellite Power System Trend Prediction Results

To validate the performance of the improved Transformer-GRU (TF-GRU) model in predicting the health status of satellite power systems, experiments were conducted using voltage and temperature data from the north energy storage battery pack of a specific type of satellite. The dataset was chronologically partitioned into training and testing sets, with the testing set containing 4490 sampling points from a complete charge/discharge cycle. The prediction results of the improved TF-GRU model, conventional Transformer model, and standard TF-GRU model are compared in Figure 2 and Table 1.

Table 1. Performance comparison of different prediction models for voltage and temperature in satellite power systems.

Parameter	Model	MAPE	RMSE
	Transformer	0.7776%	0.7179
Voltage	Normal TF-GRU	0.6308%	0.6108
	Improved TF-GRU	0.1712%	0.1729
	Transformer	1.8406%	0.4142
Temperature	Normal TF-GRU	1.8780%	0.4352
	Improved TF-GRU	1.1410%	0.2954

Figure 2. Prediction results of three models for (**a**) the north battery voltage and (**b**) the north battery temperature of a certain type of satellite.

3.1.1. Voltage Prediction Performance

As shown in Figure 2a, the improved TF-GRU model demonstrates higher accuracy in voltage trend prediction than the baseline models. While the conventional Transformer model and standard TF-GRU model exhibit significant deviations from actual voltage values in the later stages of prediction, the improved TF-GRU model closely aligns with the true data. The quantitative evaluation metrics (Table 1) indicate that the improved TF-GRU achieves an MAPE of 0.1712% and an RMSE of 0.1729 V, significantly outperforming the conventional Transformer model (MAPE = 0.7776%, RMSE = 0.7179 V) and standard TF-GRU model (MAPE = 0.6308%, RMSE = 0.6108 V). The voltage prediction results of the improved model and baseline models are subjected to a Mann–Whitney U test. The statistical analysis reveals a highly significant difference in voltage prediction errors between the improved model and the conventional Transformer architecture (Mann–Whitney U test, $p = 0.00199 < 0.01$), indicating that the performance enhancement of the proposed model is statistically validated and not attributable to random chance. This improvement is attributed to the integration of the GAT in the Transformer encoder, which dynamically adjusts attention weights to capture spatial correlations among battery cells, and the replacement of the Transformer decoder with the GRU network, which enhances temporal modeling efficiency.

3.1.2. Temperature Prediction Performance

For temperature prediction (Figure 2b), the improved TF-GRU model also exhibits robust performance. The model accurately captures periodic temperature fluctuations caused by the satellite's umbra periods (Earth shadow cycles). In contrast, the conventional Transformer model and standard TF-GRU model fail to effectively predict minimum temperature values. The quantitative results (Table 1) show that the improved TF-GRU achieves an MAPE of 1.1410% and an RMSE of 0.2954 °C, outperforming the conventional Transformer model (MAPE = 1.8406%, RMSE = 0.4142 °C) and standard TF-GRU model (MAPE = 1.8780%, RMSE = 0.4352 °C). The enhanced performance is primarily due to the GRU's ability to model sequential dependencies efficiently and the GAT's capability to refine spatial feature extraction.

3.2. Satellite Power System State Detection Results

To validate the effectiveness of the proposed model in detecting faults (e.g., short-circuit/open-circuit faults in battery cells or sudden failures in power control devices), experiments were conducted using fault data from the north energy storage battery pack. A simulated scenario was designed where the Battery Charge Regulator (BCR) fails to

disconnect after full charging, leading to persistent constant-current (CC) charging and abnormal increases in voltage (Figure 3) and temperature.

(a)

(b)
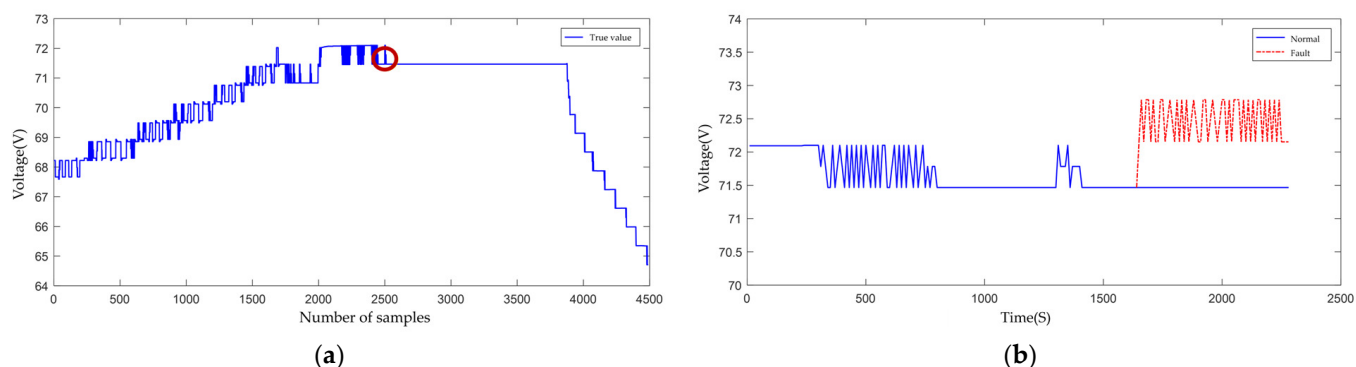
Figure 3. Schematic diagram of the fault occurrence time for the north battery pack of a certain type of satellite. In (**a**), the red circle indicates the moment when the fault occurred. Panel (**b**) shows the detailed normal and faulty voltage variation curves at the circled location.

3.2.1. Fault Detection Accuracy

As illustrated in Figure 4a, a fault occurs when the BCR remains in CC mode despite the battery pack being fully charged, resulting in a continuous voltage rise. The improved TF-GRU model successfully identifies this anomaly through a residual analysis (Figure 4b). When the predicted voltage deviation exceeds a threshold of 0.7 V, the model triggers a fault alarm. The detection residuals (Figure 4b) exhibit a sharp increase at the fault time and remain elevated for an extended period, indicating the model's ability to detect abnormal voltage changes and infer BCR malfunctions. In contrast, baseline models fail to capture the abrupt voltage rise, highlighting the limitations of traditional approaches in real-time fault detection.

(a)

(b)
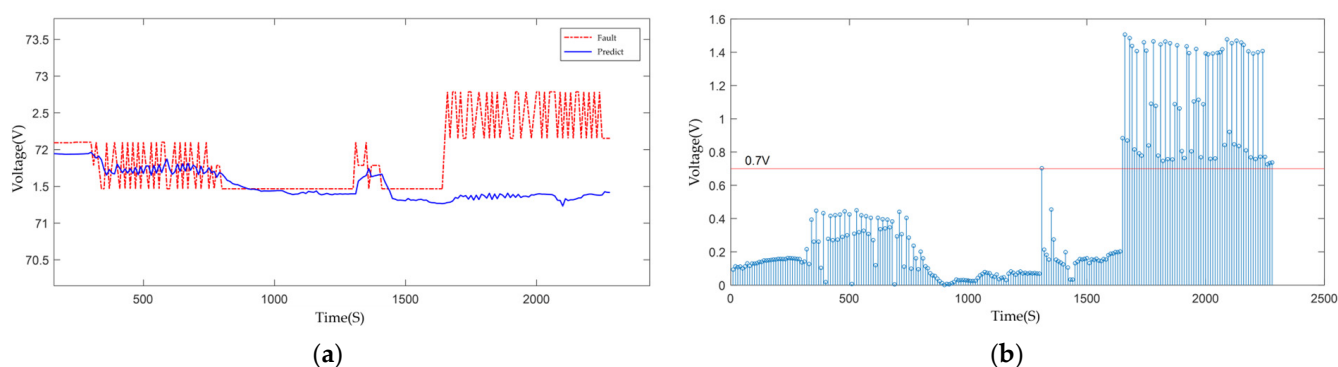
Figure 4. Panel (**a**) shows the voltage prediction and actual fault values from the improved TF-GRU model when the battery pack experienced a fault. Panel (**b**) illustrates the residual variation (blue line) detected by the improved TF-GRU model during the fault event. The voltage fault threshold is set at 0.7 V; when the residual exceeds this baseline, it indicates that a fault has occurred in the battery pack.

3.2.2. Model Robustness and Engineering Implications

The experimental results demonstrate that the improved TF-GRU model not only achieves high accuracy in trend prediction but also exhibits robustness in fault detection tasks. The model achieves rapid fault onset detection, with a false alarm rate of 1.54%. By integrating a GAT and a GRU, the model effectively integrates spatial and temporal feature extraction capabilities, enabling it to handle complex interactions among battery cells and dynamic system behaviors. By dynamically adapting the adjacency matrix based

on the schematics of different satellite batteries, the framework achieves generalization across satellite types and battery architectures. This hybrid architecture addresses the limitations of conventional methods, such as the inability of the Transformer model to capture spatial dependencies and the inefficiency of the standard TF-GRU model in handling long-term sequences.

The proposed model holds significant engineering value for satellite power management systems. It can be deployed in on-orbit monitoring systems to provide real-time health status updates and early fault warnings, thereby enhancing the reliability and safety of satellite operations. Additionally, the modular design of the model facilitates seamless integration with existing telemetry data pipelines, making it adaptable to various satellite platforms.

4. Discussion

The proposed improved TF-GRU model demonstrates several advantages in satellite power system analysis. First, integrating a GAT into the Transformer encoder represents a pioneering attempt in the field of satellite power systems. This approach first explicitly models the physical connections among battery cells through graph structures, ensuring that attention weights align with the physical laws governing fault propagation. The design transcends the limitations of conventional Transformers, which only capture implicit dependencies within sequential data, by dynamically learning the spatial correlation weights between nodes. This enables the explicit capture of spatial dependencies and localized fault propagation characteristics, achieving distributed fault localization in satellite batteries. Such a capability is critical for identifying anomalies in distributed energy storage systems. By leveraging the three-dimensional output characteristics of the model (voltage, temperature, and battery pack state), the attention weights from the GAT layers can be utilized to identify the most influential battery cells contributing to voltage/temperature variations. For instance, when voltage anomalies occur, the top three battery cells ranked by attention weights in the GAT layers are inferred as anomalous units. By correlating their physical locations, this establishes an attribution chain linking voltage anomalies to critical cells and their physical connections. Second, the adoption of a GRU-based decoder significantly reduces computational complexity compared to that of the standard Transformer decoder. This optimization ensures real-time performance requirements for satellite applications, where low latency is essential for mission-critical operations. Furthermore, the model exhibits robust fault detection capabilities by accurately identifying subtle anomalies in voltage and temperature data. This feature is particularly valuable in harsh space environments, where the early detection of faults (e.g., voltage surges or thermal deviations) can prevent cascading failures and enhance system reliability. The combination of spatial and temporal modeling through the GAT and GRU provides a comprehensive framework for addressing both trend prediction and fault diagnosis tasks.

Despite its promising performance, the improved TF-GRU model has certain limitations that warrant further investigation. One key challenge lies in the scarcity of satellite telemetry data, as in-orbit testing is costly and time-consuming. To address this, future work could explore data augmentation techniques or transfer learning strategies to improve model generalization across diverse operational scenarios. Another limitation pertains to the deployment of the model on resource-constrained satellite hardware, such as field-programmable gate arrays (FPGAs) or application-specific integrated circuits (ASICs). While the GRU-based decoder reduces computational overhead, further optimizations in model compression or quantization may be necessary to meet the stringent power and memory constraints of space platforms. Additionally, the application of the GAT in satellite battery graph modeling faces generalization challenges under sparse connectivity.

Although this study mitigates overfitting risks through a physics-constrained adjacency matrix, future research must address two critical issues: (1) dynamic variations in connection density caused by enhanced thermal coupling under extreme operational conditions and (2) graph structural drift induced by battery aging during long-term missions. To enhance model robustness, we plan to incorporate DropEdge regularization and cross-topology transfer learning in subsequent work. Multimodal data integration presents unique synchronization challenges inherent to satellite platforms: sampling frequency discrepancies and communication delays across sensors may lead to feature misalignment. To address these, this study proposes the following optimizations: (1) a sliding window interpolation-based synchronization method to sample internal resistance data to 10 Hz and (2) the introduction of cross-modal temporal attention mechanisms within the GAT layers to dynamically learn sensor-specific delay patterns. These strategies complement the temporal modeling capabilities of the existing TF-GRU architecture and establish a foundation for future engineering applications. The current study does not encompass extreme scenarios where parameters exceed the training range; however, the model design exhibits latent adaptability: physics-informed graph constraints in the GAT and temporal pattern memory in the GRU may partially generalize to anomalous operating conditions. Future work will systematically validate the model's performance in extreme cases, such as overcharging (voltage > 75 V) and deep discharge (voltage < 60 V), through physics-based data augmentation and transfer learning. This is critical for reliability assessment in satellite systems with long operational lifespans (15 years).

This study contributes to the advancement of data-driven fault detection methodologies in aerospace engineering. By integrating graph-based spatial modeling with a recurrent temporal analysis, the proposed framework offers a scalable solution for monitoring complex satellite subsystems. The approach is not limited to power systems and can be extended to other domains, such as thermal management, propulsion control, or structural health monitoring. Beyond aerospace applications, the methodology has potential implications for terrestrial systems, including electric vehicle battery management and renewable energy storage networks, where simultaneous spatial and temporal dependencies are prevalent. Furthermore, this work bridges the gap between deep learning and space engineering, demonstrating how hybrid architectures can address the unique challenges of autonomous satellite operations. By enabling real-time health monitoring and proactive fault diagnosis, the proposed model lays the foundation for more intelligent and self-sustaining spacecraft systems in future missions.

5. Conclusions

The experimental results validate the effectiveness of the improved Transformer-GRU model in both trend prediction and fault detection tasks for satellite power systems. The experimental results demonstrate the model's superior performance in both gradual parameter variation and abrupt change scenarios. For slow parameter fluctuations during normal charge/discharge processes, such as voltage plateau variations and periodic temperature oscillations, the integration of the GAT and GRU enables precise prediction through spatial correlation modeling and long-term temporal memory. In abrupt fault scenarios (e.g., short circuits and regulator failures), the model rapidly identifies anomalies by dynamically adjusting attention weights and leveraging gated mechanisms, achieving a detection delay of less than 2 s. This dual-capability characteristic renders the model applicable to full operational condition monitoring in satellite power systems. Statistical validation confirms the model's superiority in voltage prediction with highly significant results ($p < 0.002$), providing quantitative support for its reliability in engineering applications. By integrating the GAT with the GRU, the model achieves state-of-the-art performance in spatiotemporal

dependency modeling and real-time anomaly identification. By continuously capturing subtle parameter drifts, the model enables early anomaly detection prior to the activation of conventional threshold-based alarms, thereby triggering proactive warnings. Compared to traditional tools relying solely on static parameter thresholds, this approach achieves earlier identification of progressive faults. The model's unique value lies in the deep integration of trend analysis and fault diagnosis, facilitating a paradigm shift from reactive alarm systems to proactive prediction in satellite battery lifecycle monitoring. This framework is particularly advantageous for scenarios where early-stage fault signatures are subtle and prone to being overlooked by conventional methods. This study provides a reliable solution for enhancing the reliability of satellite power management systems and lays the foundation for future research on autonomous spacecraft health monitoring.

Author Contributions: Conceptualization, G.X. and J.Z.; methodology, G.X., J.Z. and Z.H.; software, G.X.; validation, G.X. and J.Z.; formal analysis, G.X. and Z.H.; investigation, J.Z. and Z.H.; resources, G.X., X.Y. and Z.H.; data curation, G.X. and J.Z.; writing—original draft preparation, G.X.; writing—review and editing, J.Z. and Z.H.; visualization, X.Y.; supervision, X.Y.; project administration, G.X., X.Y. and J.Z.; funding acquisition, G.X., X.Y. and Z.H. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest. The authors declare that there is no commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. Li, W.C. Research on On-Orbit Condition Monitoring Method for Key Devices of Satellite Power System. Master's Thesis, Nanjing University of Science and Technology, Nanjing, China, December 2020. (In Chinese).
2. Deng, Z.; Lin, X.; Cai, J.; Hu, X. Battery health estimation with degradation pattern recognition and transfer learning. *J. Power Sources* **2022**, *525*, 231110. [CrossRef]
3. Wei, J.H.; Wang, J.Q.; Mu, J.J.; He, Z.M.; Zhou, X.Y. Anomaly Detection for Satellite Power System Based on Gaussian Mixture Model. *J. Aerosp. Control Appl.* **2022**, *48*, 104–114. (In Chinese)
4. Zhang, Q.; Shang, Y.L.; Li, Y.; Cui, N.X.; Duan, B.; Zhang, C.H. A novel fractional variable-order equivalent circuit model and parameter identification of electric vehicle Li-ion batteries. *ISA Trans.* **2019**, *93*, 351–364. [CrossRef] [PubMed]
5. Hinz, H. Comparison of Lithium-Ion Battery Models for Simulating Storage Systems in Distributed Power Generation. *Inventions* **2019**, *4*, 41. [CrossRef]
6. Li, J.; Adewuyi, K.; Lotfi, N.; Landers, R.G.; Park, J. A single particle model with chemical/mechanical degradation physics for lithium ion battery State of Health (SOH) estimation. *Appl. Energy* **2018**, *229*, 692–705. [CrossRef]
7. Ansari, S.; Ayob, A.; Lipu, M.S.H.; Hussain, A.; Saad, M.H.M. Particle swarm optimized data-driven model for remaining useful life prediction of lithium-ion batteries by systematic sampling. *J. Energy Storage* **2022**, *56*, 105917. [CrossRef]
8. Fan, Y.; Xiao, F.; Li, C.; Yang, G.; Tang, X. A novel deep learning framework for state of health estimation of lithium-ion battery. *J. Energy Storage* **2022**, *56*, 105916. [CrossRef]
9. Li, W.; Fan, Y.; Ringbeck, F.; Jost, D.; Han, X.; Ouyang, M.; Sauer, D.U. Electrochemical model-based state estimation for lithium-ion batteries with adaptive unscented Kalman filter. *J. Power Sources* **2020**, *476*, 228681. [CrossRef]
10. Zhang, S.; Li, D.; Zhang, J.; Liu, G. Equivalent time-varying internal resistance model of lithium-ion batteries and the corresponding applications. *Chin. J. Sci. Instrum.* **2024**, *45*, 118–128. [CrossRef]
11. Li, Y.Q.; Lei, M.J.; Liu, P.P.; Wang, R.X.; Xu, M.Q. A Novel Framework for Anomaly Detection for Satellite Momentum Wheel Based on Optimized SVM and Huffman-Multi-Scale Entropy. *Entropy* **2021**, *23*, 1062. [CrossRef] [PubMed]
12. Khumprom, P.; Yodo, N. A Data-Driven Predictive Prognostic Model for Lithium-ion Batteries based on a Deep Learning Algorithm. *Electronics* **2019**, *12*, 660. [CrossRef]

13. Huang, Y.F.; Liang, H.C.; Xu, L. Kalman filter optimize Transformer method for state of health prediction on lithium-ion battery. *Energy Storage Sci. Technol.* **2024**, *13*, 2791–2802.

14. Song, Y.; Liu, D.; Yang, C.; Peng, Y. Data-driven hybrid remaining useful life estimation approach for spacecraft lithium-ion battery. *Microelectron. Reliab.* **2017**, *75*, 142–153. [CrossRef]

15. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.

16. Gao, Y. Research on Fault Simulation and Diagnosis of GEO Satellite Power System. Master's Thesis, Harbin Institute of Technology, Harbin, China, June 2020. (In Chinese).

17. Cao, C.; Sun, Z.; Lv, Q.; Min, L.; Zhang, Y. VS-TransGRU: A Novel Transformer-GRU-Based Framework Enhanced by Visual-Semantic Fusion for Egocentric Action Anticipation. *IEEE Trans. Circuits Syst. Video Technol.* **2024**, *34*, 11605–11618. [CrossRef]

18. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph Attention Networks. In Proceedings of the 6th International Conference on Learning Representations (ICLR 2018), Vancouver, BC, Canada, 30 April–3 May 2018.

19. Sriramulu, A.; Fourrier, N.; Bergmeir, C. Adaptive Dependency Learning Graph Neural Networks. *Pattern Recognit.* **2023**, *138*, 109348. [CrossRef]

20. Cho, K.; van Merrienboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *arXiv* **2014**, arXiv:1406.1078.