Gated Tri-Tower Transformer (GT3) - An Inflated Power Generation Attack Detector for Microgrids

Ting-Yu Syu*, Ted T. Kuo*, and Chia-Yu Lin†
*College of Artificial Intelligence, National Yang Ming Chiao Tung University, Hsinchu, Taiwan
†Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan
tysyu.ai09@nycu.edu.tw, sallylin0121@csie.ncu.edu.tw, tedkuo@nycu.edu.tw

2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC) | 979-8-3503-1019-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICBC56567.2023.10174940

***Abstract*—Renewable energy microgrids are flourishing due to the rising environmental consciousness. Blockchain technology is considered a promising solution for advanced metering infrastructures (AMIs) connecting distributed microgrids, enabling power tracking and trading among participants. However, AMIs are vulnerable to attacks, particularly the Inflated Power Generation (IPG) attack, in which energy producers manipulate smart meters to report more energy than they actually produce. Currently, detecting IPG attacks relies mainly on human intervention, which is costly and time-consuming, making the process inefficient a nd n ot s calable. A lthough b lockchain-based AMIs possess desired features, they are not immune to IPG attacks since there is no on-chain ground truth to assess the reasonability of the reported amount of energy produced. This study proposes a novel Gated Tri-Tower Transformer (GT3) neural network architecture for a multivariate time series classifier t o serve as an IPG attack detector for each AMI blockchain node to validate energy transactions. The GT3 model captures temporal correlations and fuses different features to enhance the accuracy of IPG attack detection. Our experiments with the CAISO dataset demonstrate that the proposed detector outperforms the most advanced detectors, with an increased detection rate from 71.8% to 80.5% and a reduced false alarm rate from 12.2% to 5.6%. Moreover, we investigated and concluded that GT3 performs better than prior works even during the ramp-up periods for as little as one month of collected data.**

***Index Terms*—inflated power generation attack, gated tri-tower transformer, advanced metering infrastructure, microgrid**

I. Introduction

As microgrids become increasingly popular, more individual and organizational participants are setting up microgrids and becoming renewable energy producers. The producers can trade extra energy with grid operators or other participants; and earn green certificates b ased o n t he a mount o f renewable energy generated. Blockchain technology is considered a foundational technology to build an autonomous, trustable, and robust renewable energy tracking and trading platform reducing trading frictions among participants in advanced metering infrastructures (AMIs) [1]–[4].

AMIs are an appealing target for attacks because of the potential for substantial gains and the significant n umber of participants with varied interests. Among the types of attacks that can occur in AMIs, Inflated P ower G eneration (IPG) attacks [5], [6] are launched by energy producers, while consumers carry out Electricity Theft (ET) [7] attacks.

Although blockchain technology can establish a trusted network, relying on stakeholders' smart meters to report the energy consumption or production at each site may not be the most reliable source of information. As microgrids continue to grow, there is a need for an IPG attack detector in AMI blockchain nodes that can automatically verify each energy generation report. Identifying IPG attacks currently requires expensive and time-consuming human intervention. While significant research has been conducted to address consumption-side ET attacks in the centralized power generation setting, there is a lack of efforts to address IPG attacks in the emerging decentralized power generation paradigm. In previous works [8]–[10], recursive neural networks (RNN) and autoencoders were utilized to address this issue, but their performance was limited. On the other hand, Transformers [11] have achieved remarkable success in dealing with sequence-to-sequence data with long-term dependencies, such as natural language processing and computer vision tasks. However, little research is investigating the feasibility of utilizing Transformers to detect AMI attacks. This study examines the possibility and presents a novel *Gated Tri-Tower Transformer* (GT3) neural network architecture-based IPG detector. Our study shows that when tested against the California Independent System Operator (CAISO) dataset, the proposed GT3 IPG detector outperforms previous work.

The main contributions of this study are as follows:

1) Investigation of the applicability of Transformers for detecting IPG attacks in AMIs.
2) Proposed a GT3 IPG detector based on a multivariate time series classifier that captures spatial, single-step, and multi-step temporal characteristics, which outperforms previous detectors, with an increased detection rate from 71.8% to 80.5% and a reduced false alarm rate from 12.2% to 5.6%.
3) Investigation of the effect of reduced time scales on detector training, demonstrating that the proposed detector can achieve state-of-the-art performance with only one month of training data.

The rest of this paper is organized as follows: Section II reviews prior work on IPG and ET attack detection. Section III explains the design of the proposed GT3 IPG detector, and a set of experiments and discussions are given in Section IV and the study concludes with some future work in Section V.

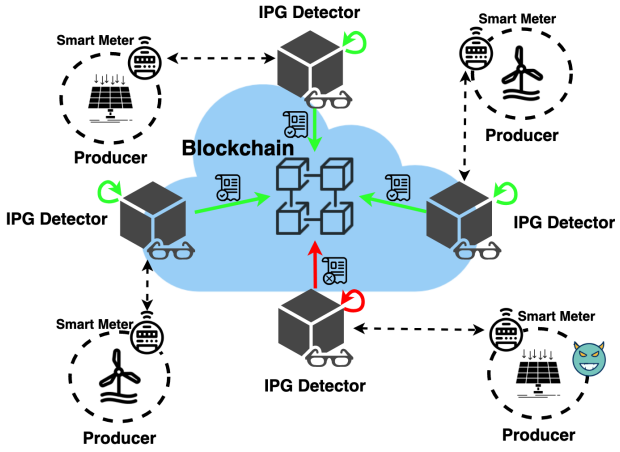

Fig. 1. Smart meters submit reports on energy produced in transactions. The ingress blockchain node performs transaction validation with the help of the IPG Detector.

II. Related Works

Although IPG attacks are growing problems in AMIs, limited research has focused on IPG attack detection. Recursive neural networks (RNN)-based solutions have been adopted to address IPG detection challenges. For instance, [7] proposed a gated recurrent unit (GRU)-based RNN classifier, and [6] developed a convolutional-recurrent neural network (CRNN) IPG attack detector. However, RNN-based methods are limited in dealing with data with long dependencies and lack parallelism, which affects their effectiveness and efficiency [12], [13]. The Transformer [11] has shown excellent results in time series data. It has been further applied in various fields, including the Gated Transformer Network (GTN) proposed in [14] and the Gated Three-Tower Transformer (GT^3) in [15]. While Transformer-based models can achieve outstanding performance in time series tasks, they have not been explored for detecting IPG attacks. This is because the self-attention layer of the traditional Transformer model can only capture the pattern between single data points [16], and energy generation data lacks additional textual information. Therefore, this study proposes an IPG attack detector using a novel *Gated Tri-Tower Transformer* (*GT3*) neural network architecture to address the challenges in detecting IPG attacks.

III. Methods

The main objective of an IPG attack detector is to validate reported power generation by producers. Renewable energy generation data being time-series data, detecting malicious attacks using meter readings and other input features is possible. To capture spatial and temporal information, GTN [14] used two encoder towers to model channel-wise and step-wise relationships and used a gate to combine the two outputs to determine whether channel-wise or step-wise relationships should be emphasized. However, the Self-Attention layer in the vanilla Transformer can only extract the similarity of a single time step [16] similar to the convolution with a kernel size equal to one, as shown in Fig. 2. A method is required to

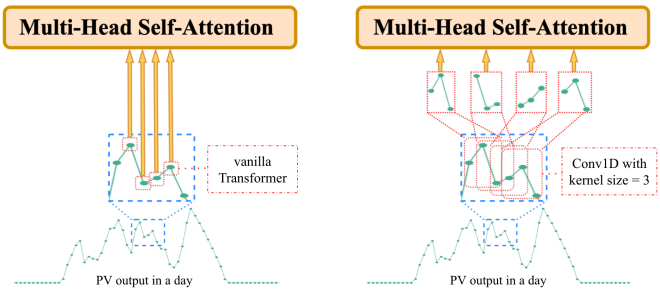

Fig. 2. The Self-Attention layer in the vanilla Transformer is similar to the convolution with a kernel size equal to 1 [16]. We need a method to extract the correlation between multiple time steps.

extract the correlation between a single time step and multiple time steps. It can only extract the similarity of a single time step. Nevertheless, the pattern in the time series may be a point or a combination of changes over time, but the Self-Attention layer is calculated based on the point-wise value of the vector. Therefore, we need a method that can extract the correlation between a single time step and the correlation between multiple time steps.

Inspired by GTN, we proposed to add a **Multi-Step Encoder** to GTN, as shown in Fig. 3, and called it a **Gated Tri-Tower Transformer** (**GT3**). Each Multi-Step Encoder block comprises multi-head self-attention (MSA) and fully connected (FC) layers. Also, each of these two layers connects to a residual connection, followed by a layer normalization(LN). However, unlike the Time-Step Encoder in GTN, we want to simultaneously compare the attention weight among multiple data points. Therefore, before entering the encoder, we use a one-dimensional convolutional (Conv1D) layer and a filter of size f to convert all data points over f periods. Compared with the Self-Attention layer, which only considers a single data point, we expect to summarize the presence of detected features in the input, capturing additional temporal correlation. Therefore, the first block can be represented as:

$$\begin{aligned} Z^{multi,1} &= LN(MMSA(Conv1D(X, f)) + X) \\ X^{multi,1} &= LN(FC(Z^{multi,1}) + Z^{multi,1}) \end{aligned} \tag{1}$$

where X is the input sequence denoted by a set of time points $x_1, x_2, ...x_L$, L is the length of the input sequence, and x_t is the data point of time t. $Conv1D$ comprises a one-dimensional convolution layer, and f is the filter (kernel) size, which controls the number of time steps in each read of the input sequence. FC is made up of two linear layers with a ReLU activation. $X^{multi,1}$ is the output of the first Multi-Step Encoder block, $Z^{multi,1}$ is the hidden representation of the first Multi-Step Encoder block, and $MMSA$ is the main component of the vanilla Transformer [11]. A masked multi-head self-attention (MMSA) is used to extract the temporal features to ensure that each time series data point can only rely on the previous features.

The proposed GT3 neural network architecture incorporates this Multi-Step Encoder, as shown in Fig. 3 to capture addi-

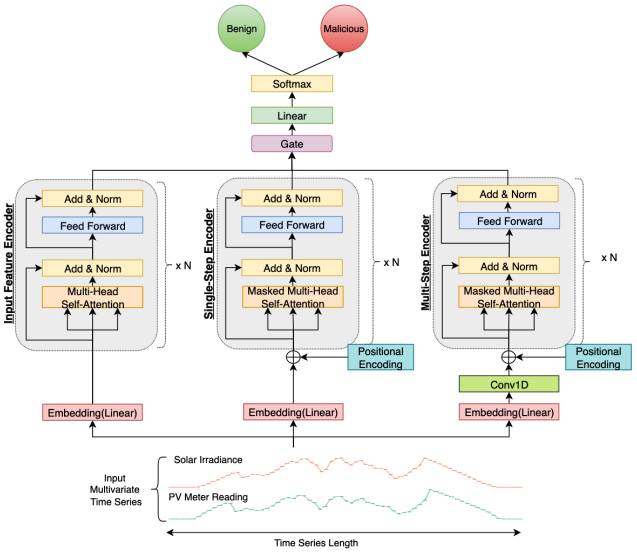

Fig. 3. The Proposed Gated Tri-Tower Transformer (GT3) Model.

tional local environments for better IPG attack detection. In addition, while retaining the Single-Step Encoder, the adaptive gating component can fuse different features to detect attacks on generation records.

A. Embedding Component

The vanilla Transformer is a sequence-to-sequence model. The input word must first be processed through the embedding layer to transform the words into vectors [11]. In the case of power generation data, since the generation is continuous, we pass it to an FC layer to get the embedding vectors with D dimensions. The two time-step encoder towers add positional encoding to specify the time steps further. The positional encoding technique used here is the same as the vanilla Transformer.

B. Encoding Component

The Encoding Component consists of three encoders: the Input Feature Encoder, the Single-Step Encoder, and the Multi-Step Encoder. The Input Feature Encoder comprises N identical encoder blocks, each comprising multi-head self-attention (MSA) and fully connected (FC) layers. To capture their correlation, the Input Feature Encoder calculates the attention weights among distinct input features in all time steps. Similarly, the Single-Step Encoder comprises N identical encoder blocks, and an MMSA is used to extract temporal features. The Single-Step Encoder calculates the attention weights between single data points in all time steps to capture temporal correlation. The Multi-Step Encoder uses a 1-dimensional convolutional (Conv1D) layer and a filter of size f to convert all data points into considering f periods before entering the encoder. This preprocessing allows the Multi-Step Encoder to summarize the presence of detected features in the input and capture additional temporal correlation among multiple data points.

C. Gating Component

Instead of simply concatenating the features of the three encoder towers, we use the adaptive Gating mechanism [14] to learn the features of each of the three towers. We set the output of the three encoder towers as $X^{feature}, X^{single}, X^{multi}$ and concatenate them into a vector. Similarly, we then fed this vector to an FC layer, and then we obtained the gating weight of each tower by using the $softmax$ activation function as follows:

$$g^{feature}, g^{single}, g^{multi} = Softmax(FC(Concat(X^{feature}, X^{single}, X^{multi}))) \tag{2}$$

Then we get the final weighted vector Y through an FC layer as follows:

$$Y = FC(Concat(g^{feature} \cdot X^{feature}, g^{single} \cdot X^{single}, g^{multi} \cdot X^{multi})) \tag{3}$$

Finally, the final inflated power generation probability vector is then obtained.

IV. Evaluation

A. Experiments Design

The recommended training time scale in previous works [6], [17] is 1 to 2 years. However, having a long warm-up period for microgrids could be an issue in terms of cost and difficulty. Therefore, we aim to evaluate the performance of the proposed GT3 model using a minimal amount of data, as little as one month, too. Previous works have not explored the effects of ramping up the training data effects as time progresses. Therefore, we design two types of experiments, *two-year* and *one-month*, for the IPG attack detector training to explore the ramp-up effects on the detectors.

Therefore, we designed two types of experiments, *two-year* and *one-month*, for the IPG attack detector training. To our knowledge, the effect of reducing the training data time scale has not been explored in previous works.

B. Experiments Settings

1) Dataset

We utilized the open power system dataset [18] to study renewable energy generation across four California Independent System Operator (CAISO) zones. We extracted the PV generation and solar irradiance per 15 minutes in these areas for three years. Due to privacy concerns, much of the malicious data in this study was generated artificially using the same four cyberattack functions proposed in [6], as listed in Table I. We balanced the data using benign data and randomly selected benign data for generating malicious data. All datasets were standardized and split into training and testing sets.

2) Models And Hyperparameters

To assess the effectiveness of the GT3, we compared it to two prior works: CRNN and GTN, using the same hyperparameters and optimizer specified in their respective literature. In addition, for comparison, we also compared a two-tower transformer-based model we developed, *mGTN*,

TABLE I
CYBERATTACK FUNCTIONS PROPOSED IN [6]

Attack Type	Representation
Partial Increment Attack	$f_1(G_{t,\ d}) = (1+\alpha)G_{t,\ d}$
Random Partial Increment Attack	$f_2(G_{t,d}) = (1+\alpha)G_{t,d}, \alpha \in random(h,l)$
Minimum Generation Attack	$f_3(min(G_{t,d})) = \alpha(max(G_{t,d}))$
Peak Generation Attack	$f_4(G_{t,d}) = max(G_{t,d}, G_{t+1,d})$
Note: $G_{t,d}$ denotes the actual power generated at time t for each day d.	

which replaced the single-step encoder in GTN with the proposed multi-step encoder. Moreover, we set the embedding dimension and hidden layers to 512, used 8 multi-heads, 4 layers of the encoder tower, and 0.2 dropout probabilities. Moreover, during training, we used adagrad optimizer with a learning rate of 1e-4, batch size of 128, and 200 epochs.

C. Experiments Results

For brevity, we present the results of Zone 1 only. The performance metrics and the average gating weight of each encoder tower are listed in Tables II and III, respectively. During the two-year experiment, GT3 outperformed other models, achieving a detection rate of 80.5% and a false alarm rate of 5.6%. Based on the experimental results, we observed that the Input Feature Tower was given more weight. The Multi-Step Tower of mGTN showed a slight increase in weight assignment compared to the Single-Step Tower of GTN, but the difference was insignificant. In contrast, in GT3, the weights of both time-step towers were distributed relatively evenly, resulting in improved detection performance. This may be because the model learned distinct features for both towers of temporal correlation, unlike GTN and mGTN, which only considered a single time-step tower.

The CRNN-based detector cannot be trained on newly constructed microgrids with only one month of training data, whereas the Transformer-based GTN achieves a 50.8% detection rate. Compared to GTN, the detection rate of mGTN rises from 50.8% to 58.5%; we believe this is because the convolution layer in the Multi-Step Encoder aids the detector in achieving better performance with a shorter time scale. GT3 also increases the detection rate from 58.5% to 61.2% after learning three different features through three towers. In Fig. 4, we investigate further the effect of different training data time scales. The experimental results show that all three Transformer-based models exhibit a significant difference in performance when compared to CRNN, and their performance gradually improves with an increase in the training data. This suggests that the Transformer-based model is more suitable for IPG attack detection scenarios. While the overall performance of mGTN is comparable to that of GTN, the convolutional layer enhances the model's performance in scenarios involving a shorter time scale, such as one to three months.

V. Conclusion

Although blockchain technology offers a promising solution for energy tracking and trading for AMIs, it is still susceptible to IPG attacks. Current detection methods are costly and time-consuming. This study proposes a novel Gated Tri-Tower Transformer (GT3) neural network architecture for a multivariate time series classifier to serve as an IPG attack detector for blockchain-based AMIs. The proposed detector outperforms the most advanced detectors, with an increased detection rate and reduced false alarm rate. Consequently, the proposed GT3 could improve the efficiency and scalability of IPG attack detection in blockchain-based AMIs.

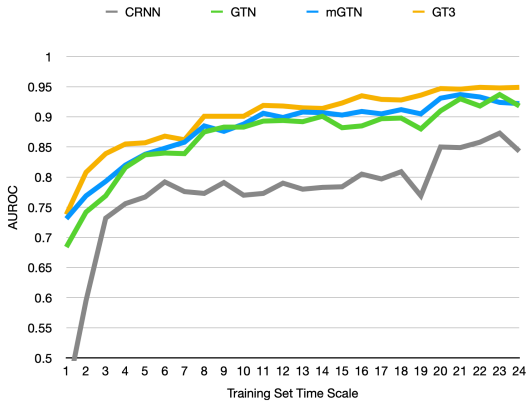

Fig. 4. The AUROC of different models with increasing training data time scales.

TABLE II
PERFORMANCE COMPARISON OF THE PROPOSED MODEL WITH CRNN AND GTN UNDER CAISO ZONE 1 FOR THE TWO-YEAR AND ONE-MONTH EXPERIMENTS

	DR	FA	PR	F1	AUROC
TWO-YEAR EXPERIMENT					
CRNN	71.8%	12.2%	84.3%	77.5%	84.3%
GTN	78.6%	9.8%	87.3%	82.7%	91.8%
mGTN	78.6%	10.0%	87.7%	82.8%	92.2%
GT3	**80.5%**	**5.6%**	**92.2%**	**85.9%**	**94.9%**
ONE-MONTH EXPERIMENT					
CRNN	-	-	-	-	43.2%
GTN	50.8%	20.0%	71.2%	59.1%	68.5%
mGTN	58.5%	**19.8%**	**74.6%**	65.3%	73.1%
GT3	**61.2%**	23.5%	71.8%	**65.7%**	**73.8.0%**
Note:	DR = detection rate = TP/(TP+FN) FA = false alarm = FP/(FP + TN) PR = precisionc= TP/(TP+FP) F1 = 2·PR·DR/(PR+DR)				

TABLE III
AVERAGE GATING WEIGHT UNDER CAISO ZONE 1 IN EXPERIMENT 1

	Input Feature Tower	Single-Step Tower	Multi-Step Tower
GTN	0.933	0.067	-
mGTN	0.916	-	0.084
GT3	0.837	0.071	0.092

REFERENCES

[1] M. B. Mollah, J. Zhao, D. Niyato, K.-Y. Lam, X. Zhang, A. M. Ghias, L. H. Koh, and L. Yang, "Blockchain for future smart grid: A comprehensive survey," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 18–43, 2020.

[2] F. Knirsch, C. Brunner, A. Unterweger, and D. Engel, "Decentralized and permission-less green energy certificates with gecko," *Energy Informatics*, vol. 3, no. 1, pp. 1–17, 2020.

[3] S. Hartnett, C. Henly, E. Hesse, T. Hildebrandt, C. Jentzch, K. Krämer, G. MacDonald, J. Morris, H. Touati, and A. Trbovich, "The energy web chain-accelerating the energy transition with an open-source, decentralized blockchain platform," *Energy Web Foundation*, 2018.

[4] J. Bao, D. He, M. Luo, and K.-K. R. Choo, "A survey of blockchain applications in the energy sector," *IEEE Systems Journal*, vol. 15, no. 3, pp. 3370–3381, 2020.

[5] Z. Yan and H. Wen, "Performance analysis of electricity theft detection for the smart grid: An overview," *IEEE Transactions on Instrumentation and Measurement*, 2021.

[6] M. Ismail, M. F. Shaaban, M. Naidu, and E. Serpedin, "Deep learning detection of electricity theft cyber-attacks in renewable distributed generation," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3428–3437, 2020.

[7] M. Nabil, M. Ismail, M. Mahmoud, M. Shahin, K. Qaraqe, and E. Serpedin, "Deep recurrent electricity theft detection in ami networks with random tuning of hyper-parameters," in *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 2018, pp. 740–745.

[8] M. E. Eddin, A. Albaseer, M. Abdallah, S. Bayhan, M. K. Qaraqe, S. Al-Kuwari, and H. Abu-Rub, "Fine-tuned rnn-based detector for electricity theft attacks in smart grid generation domain," *IEEE Open Journal of the Industrial Electronics Society*, vol. 3, pp. 733–750, 2022.

[9] A. Takiddin, M. Ismail, U. Zafar, and E. Serpedin, "Deep autoencoder-based anomaly detection of electricity theft cyberattacks in smart grids," *IEEE Systems Journal*, vol. 16, no. 3, pp. 4106–4117, 2022.

[10] N. Bhusal, M. Gautam, R. M. Shukla, M. Benidris, and S. Sengupta, "Coordinated data falsification attack detection in the domain of distributed generation using deep learning," *International Journal of Electrical Power & Energy Systems*, vol. 134, p. 107345, 2022.

[11] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[12] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[13] S. Ahmed, I. E. Nielsen, A. Tripathi, S. Siddiqui, G. Rasool, and R. P. Ramachandran, "Transformers in time-series analysis: A tutorial," *arXiv preprint arXiv:2205.01138*, 2022.

[14] M. Liu, S. Ren, S. Ma, J. Jiao, Y. Chen, Z. Wang, and W. Song, "Gated transformer networks for multivariate time series classification," *arXiv preprint arXiv:2103.14438*, 2021.

[15] J. Chen, T. Chen, M. Shen, Y. Shi, D. Wang, and X. Zhang, "Gated three-tower transformer for text-driven stock market prediction," *Multimedia Tools and Applications*, pp. 1–27, 2022.

[16] S. Li, X. Jin, Y. Xuan, X. Zhou, W. Chen, Y.-X. Wang, and X. Yan, "Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting," *Advances in neural information processing systems*, vol. 32, 2019.

[17] M. Ezeddin, A. Albaseer, M. Abdallah, S. Bayhan, M. Qaraqe, and S. Al-Kuwari, "Efficient deep learning based detector for electricity theft generation system attacks in smart grid," in *2022 3rd International Conference on Smart Grid and Renewable Energy (SGRE)*. IEEE, 2022, pp. 1–6.

[18] X. Zheng, N. Xu, L. Trinh, D. Wu, T. Huang, S. Sivaranjani, Y. Liu, and L. Xie, "Psml: A multi-scale time-series dataset for machine learning in decarbonized energy grids," *arXiv preprint arXiv:2110.06324*, 2021.