Deep Learning-Based Transmission Line Screening for Unit Commitment

Farhan Hyder Sriparvathi Shaji Bhattathiri Bing Yan Michael E. Kuhl

2023 IEEE Power & Energy Society General Meeting (PESGM) | 978-1-6654-6441-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/PESGM52003.2023.10252766

***Abstract*—Solving the transmission-constrained unit commitment (TC-UC) problem efficiently for high-quality solutions is one of the biggest challenges faced by independent system operators today. One way to achieve this is by removing superfluous network constraints from the problem. Several approaches have been used to identify such constraints. However, these methods are either too conservative or fail to maintain solution feasibility. In this paper, a deep learning-based approach is developed to identify the subset of transmission lines that can be safely removed from the TC-UC problem. The idea is to capture the temporal relationship between past line loading levels, nodal demands, and future line loading levels to identify superfluous network constraints. To achieve this, a novel regression-based classification approach is developed, where the regression model is used to predict line loading levels, and different thresholds are applied to classify transmission line capacity constraints as necessary or not for the TC-UC problem. The major advantage of this approach is that once trained, the model can be used under different classification thresholds. Numerical results show that the proposed approach significantly improves computational efficiency without compromising the solution quality.**

***Index Terms*—Transmission-constrained unit commitment, constraint screening, deep learning, LSTM**

I. INTRODUCTION

Independent system operators (ISOs) solve day-ahead TC-UC problems to determine on/off statuses and power generation levels for each unit to meet the system demand at the minimum cost [1]. The unit commitment decisions are modeled with binary variables (on (1) or off (0)) and dispatch decisions with continuous variables (generation level), making the TC-UC problem a Mixed Binary Linear Programming (MBLP) problem. The TC-UC problem consists of unit-level constraints such as generation capacity limits, ramping limits, etc., and system-level constraints such as system demand, transmission capacity limits (network constraints), etc. It has been demonstrated that the TC-UC problem is an NP-hard problem even for a single time period [2]. One approach suggested in the literature to reduce the complexity of the TC-UC problem is to filter out superfluous network constraints [3]–[9]. Many network constraints are either redundant (unloaded lines) or inactive (uncongested lines) and they can be safely removed to reduce the complexity of the TC-UC problem. However, removing too many constraints would result in infeasible solutions; whereas, removing too few constraints may not improve computational performance.

Different transmission line screening approaches are discussed in Section II. Traditional approaches often include network reduction [3], constraint generation [4] and optimization-based methods [5], [6]. While these methods maintain the feasibility of the TC-UC problem, they often fail to provide significant improvements in computational performance as they remove constraints conservatively. More recently, machine learning-based approaches have been used for transmission line screening [7]–[9]. While these methods provide significant improvements in computational performance, they might cause infeasibility in the TC-UC solution as they might remove critical constraints. In addition, the transmission line loading for future time intervals depends on nodal demand and line loading in previous time intervals. Such temporal relationships cannot be captured by the previously used machine learning-based approaches, which could be a reason for infeasibility. Recurrent Neural Networks (RNNs), which are special types of Artificial Neural Networks (ANNs) designed to work with temporal data offer a promising way to capture such non-linear temporal relationships. A simple RNN structure, however, suffers from difficulties like the vanishing and exploding gradient problems [10]. The Long Short Term Memory (LSTM), an advanced RNN architecture, overcomes the drawbacks of simple RNNs by choosing to only remember the most relevant information and performs better [11].

In this paper, a deep learning-based approach is developed to identify superfluous transmission line constraints from historical data on nodal demands and line loading levels to improve the computational performance of the TC-UC problem. In Section III, a simplified TC-UC formulation is presented. In Section IV, a novel regression-based classification approach is developed, where the regression model with LSTM layers is used to predict line loading levels and then different thresholds are applied to classify transmission line capacity constraints with the help of a rule-based system. The major advantage of this approach is that once trained, the model can be used under different classification thresholds. In Section V, the results of training and testing the model are presented along with the results of transmission line screening on a modified IEEE RTS-96 system.

Farhan Hyder and Bing Yan are with the Department of Electrical and Microelectronics Engineering, Rochester Institute of Technology, Rochester, NY, 14623, USA (email: fh6772@rit.edu; bxyeee@rit.edu). Sriparvathi Shaji Bhattathiri and Michael E. Kuhl are with the Department of Industrial and Systems Engineering, Rochester Institute of Technology, Rochester, NY, 14623, USA (email: ssb6096@rit.edu; mekeie@rit.edu). This material is based upon work supported by the National Science Foundation under Award ECCS-1810108 and DGE-2125362. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

II. Literature Review

This section reviews existing studies on transmission line screening and regression-based classification approaches.

A. Transmission line screening methods

Existing transmission line screening approaches can be divided into two categories, traditional and machine learning-based approaches. The traditional approaches focus on maintaining the feasibility of the TC-UC solution while finding the subset of transmission lines to be removed. The network reduction method shows how the size of the network can be reduced by studying the topology of the network and removing unnecessary lines and nodes [3]. The constraint generation approach iteratively adds network constraints based on necessary and sufficient conditions that maintain solution feasibility [4]. In some other approaches, the inactive constraints are identified by maximizing/minimizing the flow through a given line to determine if the line gets congested [5], [6]. These traditional approaches are often not as fast as machine leaning-based approaches.

The recent emergence of machine learning algorithms has encouraged researchers to adopt data-driven approaches toward transmission line screening. A data-driven method based on the K-nearest neighbors algorithm is used to determine if a line should be classified as congested or not for a previously unseen time interval [7]. This study shows a significant reduction in prepossessing as well as solving times. However, for highly congested and large networks, this method could give inaccurate solutions. Some recent approaches used neural networks to identify active constraint sets [8], [9]. These two approaches improved computational efficiency significantly, but they cannot guarantee correct predictions, requiring additional steps to prevent infeasible solutions. These machine learning-based approaches are faster than traditional approaches, but cannot guarantee feasible TC-UC solutions.

B. Regression-based classification and deep learning

Machine learning can be used for both classification and regression. Classification models usually predict labels for ordinal data while regression models are used to predict continuous data. For transmission line screening, both classification [7], [8] and regression [9] models have been used in the literature. If the classification of a line capacity constraint as necessary or not is based on a threshold applied to the line loading levels, the line loading levels can be predicted using a regression model and a rule-based classifier can be applied. Such models tend to learn more information as well as have the versatility of being trained once and then being used to classify based on different rules and thresholds [12].

Deep learning algorithms are emerging as promising approaches to learning non-linear relationships in power system problems [13], [14]. When input and output data have temporal relationships, RNNs are a suitable solution. RNNs remember information across time intervals using feedback loops, which connect past and current information. However, simple RNN structures offer little to no control over what information is remembered. Furthermore, simple RNNs suffer from vanishing and exploding gradient problems when back-propagating errors over time [10]. These drawbacks are overcome by advanced RNN structures called LSTMs, which are designed to offer control over memory and remember only the most relevant information while preventing gradients from vanishing or exploding [11].

III. TC-UC Formulation

This section describes the TC-UC formulation. For unit g at node n, at each time t, major decision variables include unit on/off status $x_{n,g_n,t}$ (binary) and generation level $p_{n,g_n,t}$ (continuous). Since the focus of the paper is to identify the subset of transmission lines that can be removed without impacting solution feasibility, the standard TC-UC formulation presented in [15] is simplified as follows without considering inter-temporal constraints and commitment costs following the existing work [7]:

$$\underset{p_{n,g_n,t},x_{n,g_n,t}}{\text{Min}} \sum\nolimits_t \sum\nolimits_n \sum\nolimits_{g_n} C_{g_n} p_{n,g_n,t} \tag{1}$$

$$s.t. \quad \sum\nolimits_n \sum\nolimits_{g_n} p_{n,g_n,t} = \sum\nolimits_n P^D_{n,t} \quad \forall t \tag{2}$$

$$x_{n,g_n,t} P^{\min}_{n,g_n} \leq p_{n,g_n,t} \leq x_{n,g_n,t} P^{\max}_{n,g_n} \quad \forall n, g_n, t \tag{3}$$

$$-f^{\max}_l \leq f_{l,t} \leq f^{\max}_l \quad \forall l, t \tag{4}$$

$$f_{l,t} = \sum\nolimits_n \alpha_{l,n} \left(\sum\nolimits_{g_n} p_{n,g_n,t} - P^D_{n,t} \right) \forall l, t \tag{5}$$

The objective function (1) minimizes the total generation cost. System balance condition (2) ensures that the total generation equals the total demand for every time period. Generation capacity limits (3) require that if a unit is online, its generation level $p_{n,g_n,t}$ should be within its minimum $P^{\min}_{n,g_n}$ and maximum $P^{\max}_{n,g_n}$. The transmission line capacity limits (4) ensure that the DC power flow $f_{l,t}$ of line l cannot exceed the line's capacity limit f^{max}_l. The DC power flow $f_{l,t}$ is expressed in (5) as the linear combinations of nodal injections from all nodes weighted by generation shift factor $\alpha_{l,n}$. For simplicity, it is assumed that nodal demands are known with certainty and system configuration does not change.

IV. Methodology

In this section, a deep learning-based approach developed for network constraint screening is discussed.

A. Input and output

Similar to electricity demand depending on historical data, the amount of power flowing through a transmission line at a given time interval depends on the line loading and nodal demands in the previous intervals. To understand this temporal dependence three input-output configurations are considered:

1) *X-$\hat{Y}$*: Input: Past nodal demands *X*; Output: Future line loading levels $\hat{Y}$;
2) *Y-$\hat{Y}$*: Input: Past line loading levels *Y*; Output: Future line loading levels $\hat{Y}$;

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

3) X,Y-$\hat{Y}$: Input: Past nodal demands and line loading levels X,Y; Output: Future line loading levels $\hat{Y}$.

The output $\hat{Y}$ from the LSTM-based model is used to identify the subset of transmission line capacity constraints required to solve the TC-UC problem without infeasibility occurring for any hour of the day. Therefore, the subset of required constraints changes for every hour of the day. This requires the line capacity constraints for every line l to be classified as necessary (label - 1) or not necessary (label - 0) for every time interval t. The line loading level $\tilde{f}_{l,t}$ is expressed as a ratio of the absolute value of the line flow $f_{l,t}$ to the maximum capacity of the line f_l^{max} as shown in (6) below,

$$\tilde{f}_{l,t} = \frac{|f_{l,t}|}{f_l^{\max}}. \tag{6}$$

If $\tilde{f}_{l,t}$ is below a certain threshold, the line constraint is classified as rejected (0), and above it is retained (1). Nodal demands are normalized using z-score method.

B. LSTM-based model architecture

Since the input-output data is continuous, a regression based LSTM model is used to predict the line loading and then threshold-based labeling is done for classification. The LSTM-based model developed in this study consists of an input layer (I/P), an output layer (O/P), LSTM and fully connected (FC) layers, as shown in Fig. 1. To capture the temporal relationship between input and output, the LSTM-based model looks back at a certain number of days D to predict $\hat{Y}$ for next day's TC-UC problem. So the I/P time-step = D*24hours. The I/P size also depends on input-output configuration. Whereas, the O/P size only depends on the number of lines L, O/P neurons = L*24 (e.g., for L = 120, O/P neurons = 2880). A sliding window that shifts by an hour at a time is used to generate the input-output pairs. The model parameters like the number of layers, number of neurons per layer, and activation functions are tuned for best results as shown in Fig. 1. The hyper-parameters such as learning rate, batch size, number of epochs, etc., are discussed in V-B.

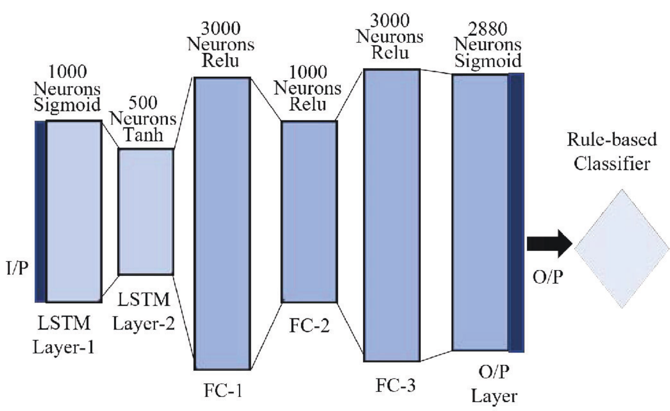

Fig. 1: Deep learning-based model for transmission line screening

The next step is to use a rule-based classifier to apply a threshold to the output from the regression model. The transmission line constraints whose $\tilde{f}_{l,t}$ is below the threshold are removed (class - 0) and above the threshold are retained (class - 1). This regression-based classification approach allows the model to be trained once and then be used to explore different classification thresholds. Mean square error (MSE) is the loss function used for training the model. Classification accuracy is used to evaluate the performance of each input-output configuration. The best-performing configuration is chosen for transmission line screening, referred to as the regression-based classification model (RCM). Different classification thresholds are then explored to achieve a better computational performance on the TC-UC problem while ensuring solution feasibility. In case line constraints are violated for any time intervals for certain classification thresholds, the violated constraints are iteratively added to TC-UC problem following the standard constraint generation (CG) procedure [4].

C. Methods for comparison

The performance of RCM is compared with other methods to evaluate its effectiveness. The baseline results are based on the network with all transmission line capacity constraints called full network (FN). Other methods include single-bus network (SBN), historical analysis (HS) and perfect prediction method (PP). These are similar to the comparative methods used in [7]. SBN ignores all network constraints and assumes a single-bus. HS retains constraints for lines that have ever been congested in the historical data and ignores the rest. PP assumes perfect forecast of transmission lines likely to get congested for any given hour and hence only retains the constraints corresponding to those lines.

V. Testing and Results

In this section, the results of transmission line screening are presented and discussed. The testing system and historical data are described in Subsection A. Then the results of training and testing the LSTM-based model are explained in Subsection B. The results of transmission line screening for the TC-UC problem are discussed in Subsection C.

A. Testing system and data

A modified IEEE RTS-96 system consisting of 96 generators, 73 nodes, and 120 transmission lines is used in this study [16], [17]. Historical nodal demands $P_{n,t}^D$ for 360 days, previously used in [7] and available at [18], are used to train the deep learning model. The simplified TC-UC problem is solved for 360 days to generate $\tilde{f}_{l,t}$. Then the historical data is split into training and testing data sets. The 360 days are first divided into four quarters of 90 days each ($Q_1 - Q_4$). Then for each quarter, 80% data (72 days) is used for training and 20% data (18 days) is used for testing. In total, 288 days are used for training and 72 days for testing. The aim of this train-test data split is to ensure that the LSTM-based model learns the seasonal variations in $P_{n,t}^D$ and $\tilde{f}_{l,t}$ over the year. All experiments are performed using IBM ILOG CPLEX Optimization Studio V 12.10.0.0 on a PC with 2.30 GHz Intel(R) Core (TM) i7- 10510U CPU and 16 GB RAM. The stop-gap (the relative difference between the objectives of the optimal relaxed solution and the current integer solution) is set to $1e-6$.

B. Training and testing results of the LSTM-based model

The results of training and testing the LSTM-based model are discussed for the three input-output configurations. To capture the weekly variations in $P^D_{n,t}$ and $\tilde{f}_{l,t}$, a look-back period of 8 days (D) is used, resulting in I/P time-step = 192 (8*24). Therefore, the LSTM-based model uses historical data from previous 8 days (e.g., Sunday of last week to Sunday of this week) to predict the output of the 9^{th} day (e.g., Monday of this week). The input-output configuration determines rest of the I/P size as follows, (1) *X-$\hat{Y}$* $\rightarrow 73*192$, (2) *Y-$\hat{Y}$* $\rightarrow 120*192$, and (3) *X,Y-$\hat{Y}$* $\rightarrow (73+120)*192$. The O/P size = $120*24$ is same for every configuration. The best hyper-parameters are a learning rate of 0.0001, and a batch size of 200 trained for 100 epochs. The optimizer used for training is Adam. The average training time for any configuration is 9.7 hours.

The MSE loss obtained from training and testing the three different input-output configurations is shown in Table I. It can be seen that the configuration *X,Y-$\hat{Y}$* gives the least MSE for training and testing.

TABLE I: THE LOSS OBTAINED FROM TRAINING THE THREE DIFFERENT INPUT-OUTPUT CONFIGURATIONS

Input-Output Configuration	Training MSE	Testing MSE
X-$\hat{Y}$	0.002995	0.005962
Y-$\hat{Y}$	0.001647	0.003302
XY-$\hat{Y}$	0.001515	0.003259

After training the regression models, the classification is done based on different thresholds (0.1-0.9). The threshold-based classification accuracy for unseen test data is shown in Table II. It can be seen that the configuration *XY-$\hat{Y}$* gives the highest accuracy for every threshold. The results in Tables I and II show that the *XY-$\hat{Y}$* input-output configuration performs best with the least mean square error and highest classification accuracy. Hence, its output is chosen for transmission line screening.

TABLE II: THRESHOLD-BASED CLASSIFICATION ACCURACY (%) FOR TEST DATA WITH DIFFERENT INPUT-OUTPUT CONFIGURATIONS

Threshold	Accuracy X to Ŷ	Accuracy Y to Ŷ	Accuracy XY to Ŷ
0.1	93.254	94.698	95.103
0.2	93.603	95.120	95.252
0.3	93.687	95.209	95.298
0.4	94.175	95.772	95.933
0.5	94.771	96.137	96.230
0.6	95.029	96.339	96.389
0.7	95.680	96.676	96.795
0.8	97.668	98.220	98.289
0.9	98.883	99.193	99.195

C. Results of transmission line screening for TC-UC

The predictions made for different classification thresholds are tested on the modified IEEE RTS-96 system using unseen testing data from each quarter (Q_1-Q_4). To explore the impact of screening on computational efficiency, the TC-UC problem is solved for the entire quarter (18 days = 432 hours) following the testing procedure in [7], resulting in 51,840 network constraints (120 constraints $*$ 432 hours) for each quarter. FN is used as a benchmark to compare the performance of RCM and other methods using the following metrics:

- Percentage Network Constraints Removed ($\%NCR$): Ratio of the number of constraints removed (NCR) to the total number of constraints (NC), $\%NCR = NCR/NC*100$;
- Solving Time (T, seconds): Time required to solve the simplified TC-UC problem;
- Percentage reduction in solving time (ΔT): Solving time of T_{RCM} relative to the solving time of T_{FN}, $\Delta T = (T_{FN}-T_{RCM})/T_{FN}*100$;
- Cost (C, \$): The total dispatch cost for the simplified TC-UC problem;
- Percentage error in cost (ΔC): Cost of C_{RCM} relative to the cost of C_{FN}, $\Delta C = (C_{FN}-C_{RCM})/C_{FN}*100$;
- Number of Violations (V): Number line constraint violations in the TC-UC solution;
- Percentage Violations ($\%V$): Ratio of the number of line constraint violations to the total number of line constraints (NC) in the TC-UC problem, $\%V = V/NC*100$.

Table III shows the testing results of each quarter (Q) where RCM with different classification thresholds (*Th*) is compared with other methods. Since FN retains all network constraints, it guarantees a feasible solution but requires a high solving time. Whereas SBN removes all constraints, resulting in over 90% reduction in solving time, but leads to significant constraint violations. PP only retains constraints corresponding to lines that are known to be congested, it significantly reduces solving time but gives infeasible solutions. By removing constraints for lines that were never congested historically, HS reduces solving time considerably. However, it results in violations for Q_1. Typically, HS performs well only for small networks with predictable line flows. RCM performs well with no infeasibility for thresholds between 0.3 and 0.6 and Th = 0.6 gives the best computational performance where the solving time is reduced by up to 59% by removing 74-75% transmission line capacity constraints. However, for higher thresholds (0.7-0.8), infeasibility occurs due to few network constraint violations, which can be resolved by iteratively adding the violated constraints back to TC-UC using CG. Given that the 0.6 threshold offers a significant reduction in solving time, there is no need to remove constraints under higher thresholds. Hence, even without the use of CG, RCM offers significant improvement in computational efficiency. For all thresholds, under all quarters, the low value ΔC implies that the solution quality obtained by RCM is very close to that of FN. These results show that the RCM offers a significant reduction in

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

solving time without causing infeasibility while maintaining high solution quality.

TABLE III: RESULTS OF TRANSMISSION LINE SCREENING: COMPARING METHODS

Q	Method	Th	NCR (%)	T (sec)	ΔT (%)	C ($)	ΔC (%*10^{-4})	V	V (%)
Q1	FN	NA	0	127	NA	31,412,161	NA	0	0
	SBN	NA	100	7	94.49	31,359,980	1661	643	1.25
	HS	NA	95	37	70.87	31,411,640	17	42	0.09
	PP	NA	99.32	7	94.49	31,403,339	281	359	0.7
	RCM	0.3	47.06	89	29.92	31,411,791	12	0	0
	RCM	0.4	57.59	78	38.58	31,411,886	9	0	0
	RCM	0.5	67.67	75	40.94	31,412,099	2	0	0
	RCM	0.6	76.33	59	53.54	31,412,211	-2	0	0
	RCM	0.7	87.28	45	64.57	31,411,208	30	4	0.01
	RCM+CG	0.7	87.28	44	65.35	31,412,346	-6	0	0
	RCM	0.8	94.41	45	64.57	31,411,628	17	18	0.03
	RCM+CG	0.8	94.41	44	65.35	31,411,988	5	0	0
Q2	FN	NA	0	127	NA	37,597,178	NA	0	0
	SBN	NA	100	8	93.71	37,552,842	1180	622	1.2
	HS	NA	94.17	35	72.45	37,597,537	-10	0	0
	PP	NA	99.14	9	92.92	37,579,851	461	385	0.75
	RCM	0.3	46.56	72	43.31	37,597,210	-1	0	0
	RCM	0.4	57.62	61	51.97	37,596,978	5	0	0
	RCM	0.5	67.5	56	55.91	37,597,044	4	0	0
	RCM	0.6	76.26	52	59.06	37,597,589	-11	0	0
	RCM	0.7	87.4	40	68.50	37,596,787	10	11	0.02
	RCM+CG	0.7	87.4	36	71.65	37,597,596	-11	0	0
	RCM	0.8	94.18	38	70.08	37,597,088	2	37	0.07
	RCM+CG	0.8	94.18	37	70.87	37,597,438	-7	0	0
Q3	FN	NA	0	83	NA	34,536,211	NA	0	0
	SBN	NA	100	8	90.37	34,485,626	1465	647	1.25
	HS	NA	94.17	34	59.04	34,535,810	12	0	0
	PP	NA	99.73	6	92.78	34,487,369	1415	634	1.23
	RCM	0.3	46.46	57	31.33	34,536,646	-13	0	0
	RCM	0.4	56.63	49	40.96	34,535,981	7	0	0
	RCM	0.5	66.7	49	40.96	34,535,910	9	0	0
	RCM	0.6	75.72	47	43.37	34,535,890	9	0	0
	RCM	0.7	86.79	39	53.01	34,535,821	11	12	0.02
	RCM+CG	0.7	86.79	37	55.42	34,536,026	5	0	0
	RCM	0.8	93.83	41	50.60	34,535,819	11	29	0.06
	RCM+CG	0.8	93.83	33	60.24	34,535,977	7	0	0
Q4	FN	NA	0	124	NA	39,857,395	NA	0	0
	SBN	NA	100	7	94.36	39,821,810	893	658	1.27
	HS	NA	92.5	38	69.36	39,857,302	3	0	0
	PP	NA	99.7	7	94.36	39,828,969	714	580	1.12
	RCM	0.3	46.37	99	20.16	39,857,681	-7	0	0
	RCM	0.4	57.05	93	25.00	39,857,821	-11	0	0
	RCM	0.5	67.44	83	33.06	39,857,806	-10	0	0
	RCM	0.6	76.39	78	37.10	39,856,804	15	0	0
	RCM	0.7	87.5	71	42.74	39,856,653	19	3	0.01
	RCM+CG	0.7	87.5	67	45.97	39,856,628	19	0	0
	RCM	0.8	94.06	69	44.35	39,856,979	10	9	0.02
	RCM+CG	0.8	94.06	63	49.19	39,856,947	11	0	0

VI. CONCLUSION

Improving the computational efficiency of the TC-UC problems while maintaining solution feasibility is of great importance for ISOs. Removing superfluous network constraints is a potential way of achieving this goal. In this paper, a deep learning-based constraint screening method is developed. To identify superfluous network constraints for the TC-UC problem, the idea is to capture the temporal relationship between past line loading levels, nodal demands, and future line loading levels. To achieve this, a novel regression-based classification approach is developed, where the regression model with LSTM layers is used to predict line loading levels and then different thresholds are applied to classify transmission line capacity constraints as necessary or not for the TC-UC problem. The major advantage of this approach is that once trained, the model can be used under different classification thresholds. Historical nodal demands and line loading levels are used for training and testing. Results of transmission line screening on the TC-UC problem show a significant reduction in solving time without compromising solution quality, demonstrating that the approach is promising. In the future, TC-UC problem with inter-temporal constraints will be considered along with large-scale networks to further demonstrate the performance of our approach.

REFERENCES

[1] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power generation, operation, and control*. John Wiley & Sons, 2013.

[2] P. Bendotti, P. Fouilhoux, and C. Rottner, "On the complexity of the unit commitment problem," *Annals of Operations Research*, vol. 274, no. 1, pp. 119–130, 2019.

[3] J. Ostrowski and J. Wang, "Network reduction in the transmission-constrained unit commitment problem," *Computers & Industrial Engineering*, vol. 63, no. 3, pp. 702–707, 2012.

[4] H. Wu, X. Guan, Q. Zhai, and H. Ye, "A systematic method for constructing feasible solution to scuc problem with analytical feasibility conditions," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 526–534, 2011.

[5] Q. Zhai, X. Guan, J. Cheng, and H. Wu, "Fast identification of inactive security constraints in scuc problems," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1946–1954, 2010.

[6] L. A. Roald and D. K. Molzahn, "Implied constraint satisfaction in power system optimization: The impacts of load variations," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 308–315, IEEE, 2019.

[7] S. Pineda, J. M. Morales, and A. Jiménez-Cordero, "Data-driven screening of network constraints for unit commitment," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3695–3705, 2020.

[8] Y. Yang, Z. Yang, J. Yu, K. Xie, and L. Jin, "Fast economic dispatch in smart grids using deep learning: An active constraint screening approach," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 11030–11040, 2020.

[9] D. Deka and S. Misra, "Learning for dc-opf: Classifying active sets using neural nets," in *2019 IEEE Milan PowerTech*, pp. 1–6, IEEE, 2019.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[11] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber, *et al.*, *Gradient flow in recurrent nets: the difficulty of learning long-term dependencies*. A field guide to dynamical recurrent neural networks. IEEE Press In, 2001.

[12] R. Adamczak, A. Porollo, and J. Meller, "Accurate prediction of solvent accessibility using neural networks–based regression," *Proteins: Structure, Function, and Bioinformatics*, vol. 56, no. 4, pp. 753–767, 2004.

[13] M. Tuo and X. Li, "Security-constrained unit commitment considering locational frequency stability in low-inertia power grids," *IEEE Transactions on Power Systems*, 2022.

[14] M. Zhou, B. Wang, and J. Watada, "Deep learning-based rolling horizon unit commitment under hybrid uncertainties," *Energy*, vol. 186, p. 115843, 2019.

[15] F. Hyder, B. Yan, M. A. Bragin, and P. B. Luh, "Impacts of uc formulation tightening on computation of convex hull prices," in *2021 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 01–05, IEEE, 2021.

[16] C. Grigg *et al.*, "The ieee reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 1010–1020, 1999.

[17] H. Pandžić, Y. Dvorkin, T. Qiu, Y. Wang, and D. S. Kirschen, "Toward cost-efficient and reliable unit commitment under uncertainty," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 970–982, 2015.

[18] OASYS, "Data IEEE 96," *GitHub repository (https://github.com/groupoasys/data_ieee96)*, 2019.

978-1-6654-6441-3/23/$31.00 ©2023 IEEE