Predicting Onset Time of Cascading Failure in Power Systems Using a Neural Network-Based Classifier

Junyuan Fang, Dong Liu, and Chi K. Tse
Department of Electrical Engineering, City University of Hong Kong, Hong Kong
Email: junyufang2-c@my.cityu.edu.hk, dongliu@cityu.edu.hk, chitse@cityu.edu.hk

***Abstract*—Cascading failure modeling and analysis provide convenient tools for assessing and enhancing the robustness of power systems against severe power outages. In this paper, we apply a neural network-based classifier to predict the onset time of cascading failure. Onset time, which has been reported as the time when the number of component failure begins to rapidly increase in the failure propagation, serves as a crucial metric to evaluate the vulnerability of power systems to cascading failure. We formulate the prediction task as a multi-class classification problem and adopt a neural network-based classifier where topological and electrical information of a power system network can be exploited for learning. Experimental results on the UIUC 150-Bus power system demonstrate a high classification accuracy by only leveraging the initial states of power networks and the initial failure sets containing the power components to be tripped at the beginning of cascading failure.**

I. Introduction

Power outages cause severe disruptions to modern society where electricity supply is essential for supporting almost all activities. In 2003 [1], a power blackout in North American due to ineffective protection strategies caused a disruption of 1800 MW of electric loads and affected around 50 million people for more than one week. According to the review reports on severe power outages in India (2012) [2] and United Kingdom (2019) [3], the inappropriate and late protection mechanism was regarded as one of the primary reasons for causing such large-area cascading failure in power systems. In recent years, power cuts occurring in Taiwan and Texas in 2021 reaffirmed the important role of prompt and precise control strategies in preventing catastrophic failure propagation. A lot of effort has been devoted to the study of cascading failure modeling and analysis in the last two decades [4]–[7] for assessing and enhancing the robustness of power systems against cascading failure.

A number of publications have appeared in recent years documenting the implementation of machine learning-based approaches for addressing challenging issues of power systems such as reducing the risk of cascading failure. In Kim *et al.*'s work [8], a graph convolutional network model is used to predict the optimal load-shedding ratio that prevents transmission lines from being overloaded under line outages. Liu *et al.* [9] extracted multiple network-based features for effectively enhancing the robustness of power networks by a decision-tree-based learning model. Other machine learning-based algorithms including support vector machine [10] and reinforcement learning [11] are applied to predict cascading fault chains and identify critical lines. As stability control is essential to mitigating cascading failure as well, studies show the effectiveness of deep reinforcement learning in intelligent voltage and frequency stability control [12], [13]. These studies clearly demonstrate that machine learning-based methods achieve outstanding performance in solving problems that are difficult to be tackled by traditional model-based methods.

In this paper, we consider the onset time based on the previous studies [9], [14] as a crucial indicator for preventing power systems from large-area cascading failure. A shorter onset time indicates that protective actions are required to be taken more timely, implying that power systems are more vulnerable to cascading failure. To effectively estimate when a certain large number of component failure events would occur in a short time during failure propagation, we adopt a neural network-based classifier to classify the onset time into three different classes in terms of the *degree of urgency.* The learning model exploits the topological information of power systems which are extracted from the initial steady state of the UIUC 150-bus power network [15], and the features obtained from comparing the magnitudes of node and line power before and after a small disturbance, i.e., the k lines that are tripped initially corresponding to the well-known $N-k$ security criteria. We perform training and testing on the samples generated from simulations, from which complete dynamic profiles of the cascading failure propagation are revealed and the onset time of each cascading failure scenario is identified. The experimental results demonstrate the effectiveness of the proposed neural network-based classifier in the dataset under various circumstances.

II. Model Description

In this section, we first introduce the onset time defined in the previous study [9], [14] and formulate the onset time prediction as a multi-class classification problem. We aim to propose a neural network-based classifier where the designated input contains the topological and electrical information of the power networks and three classes to be classified are characterized based on the length of onset time. Fig. 1 depicts a

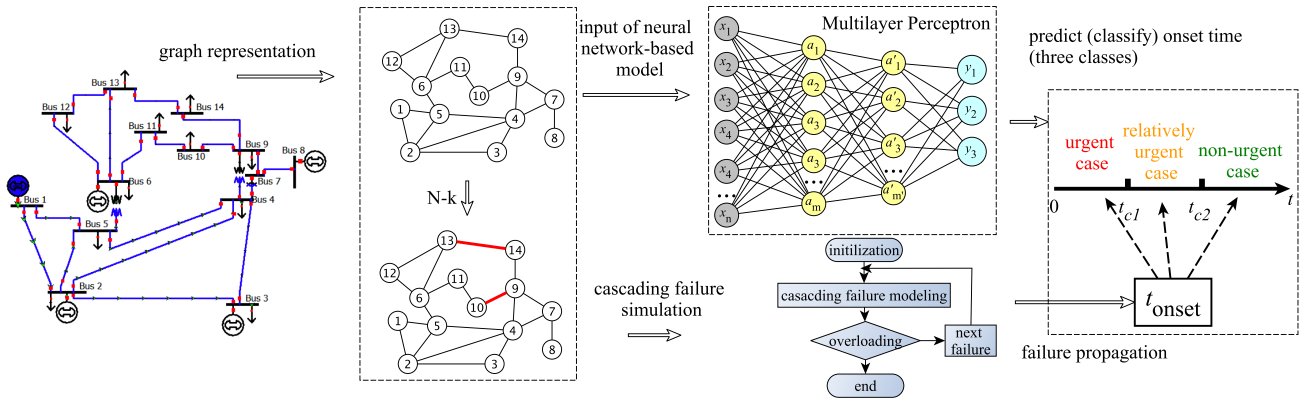

Fig. 1: A systematic framework of the application of neural network-based classifier for predicting the degree of urgency of the onset time that is identified from failure propagation profile in a power system.

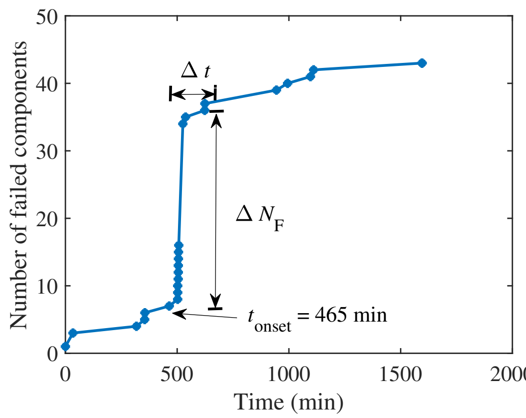

Fig. 2: Simulated failure propagation profile of UIUC 150-bus power network where two lines, i.e., from bus 2 to bus 14 and from bus 80 to bus 82, are initially tripped.

systematic framework of the application of a neural network-based classifier for predicting the degree of urgency of the onset time identified from the failure propagation profile in power systems.

A. Onset Time of Failure Propagation

In this work, we use a combined model of circuit-based power flow and stochastic processes to generate dynamic profiles of cascading failures in power systems. A typical failure propagation profile, as shown in Fig. 2, is produced by initially tripping two lines in the UIUC 150-bus test power system. From Fig. 2, the number of failed components slowly increases at the beginning of the failure propagation. At around 465 min, a number of power components fail within a short time. The onset time t_{onset}, defined as the time after which the failure propagation rate increases dramatically, can be identified when the number of failed components, denoted by ΔF during a short enough time duration, i.e., $\Delta t < \Delta t_{\text{short}}$, is larger than a given threshold ΔF_{large}. It is noted that a short onset time implies a higher vulnerability of the power system to cascading failure since protective actions are required to be taken in more timely manner.

B. Predictors of Neural Network-Based Classifier

Our main objective here is to predict the onset time of the cascading failure without generating the full profile of failure propagation which requires time-consuming simulations. In other words, our goal is to infer the onset time with the information of the initial state of the network. Instead of directly obtaining the value of onset time, we aim to implement a neural network-based classifier to predict the class of the onset time. According to the time length, we classify the onset time into three classes, namely, *urgent*, *relatively urgent*, and *non-urgent* cases, for indicating how emergent protective actions should be taken to terminate a fast and large failure propagation in a power system. To differentiate these three cases, we define two critical time points, t_{c1} and t_{c2}, and obtain three time intervals, $[0, t_{c1})$, $[t_{c1}, t_{c2})$, and $[t_{c2}, +\infty)$ accordingly. Thus, the ultimate task of the neural network-based classifier is to predict in which time interval the onset time is located. It is noted that the case when no onset time is found from a failure propagation profile belongs to the non-urgent case. The input variable, namely the predictor of the neural network-based classifier, is given by the power matrix difference ΔM_p, i.e.,

$$\Delta M_p = M_{p,N} - M_{p,N-k} \tag{1}$$

where $M_{p,N}$ is a $N \times N$ matrix obtained from the initial steady state of a power system that is represented by a N-node graph. After fixing all variables in the test power network including the admittance matrix, node injected currents and voltages, and others, the power on each power line (from bus i to bus j) is calculated by performing power flow calculation and is then entered into the element $M_{p,N}(i, j)$. The diagonal element $M_{p,N}(i,i)$ refers to the power demanded on each loading node or power supplied at each generator node. For $M_{p,N-k}$, it is the matrix obtained by calculating the power on power lines and nodes after removing k power lines and changing the admittance matrix accordingly.

Moreover, according to the definition of *power difference matrix*, the corresponding $\Delta M_p(i, j)$ equals 0 when there is no transmission line between buses i and j. Such zero-value features contribute nothing but only improve the learning of our model. Therefore, for each ΔM_p representing a specific line removal operation, we only extract the informative features, including the $\Delta M_p(i, i)$ representing the change of power demanded by loading nodes or power supplied by generation nodes, and corresponding $\Delta M_p(i, j)$ when there is a transmission line between buses i and j. Denoting the above two types of features as X_n and X_l, we can obtain the final features $X = \{X_n, X_l\}$ using a concatenation operation. The features X will be utilized as the input of our learning model.

C. Neural Network-Based Classifier

Using the power difference matrix ΔM_p obtained from (1), we develop a neural network-based classifier to predict the class of onset time. Specifically, we establish a multilayer perceptron (MLP) model [16], [17], which has two hidden layers with h_1 and h_2 hidden units to bridge the original features and the class of the onset time.

For a specific sample, we first take the original features $X \in \mathbb{R}^{1\times d}$ as inputs, where d is the dimension of features, and outputs $Y_1 \in \mathbb{R}^{1\times h_1}$ by utilizing a weight matrix $W_1 \in \mathbb{R}^{d\times h_1}$ and a bias vector $b_1 \in \mathbb{R}^{1\times h_1}$, together with a ReLU activation function, i.e.,

$$Y_1 = \mathrm{ReLU}(XW_1 + b_1). \tag{2}$$

In the next layer, we utilize a similar operation and obtain $Y_2 \in \mathbb{R}^{1\times h_2}$, but take the outputs of the previous layer as inputs, i.e.,

$$Y_2 = \mathrm{ReLU}(Y_1 W_2 + b_2), \tag{3}$$

where $W_2 \in \mathbb{R}^{h_1\times h_2}$ and $b_2 \in \mathbb{R}^{1\times h_2}$ are the weight matrix and bias vector, respectively.

Finally, we obtain the final outputs as

$$Y_3 = Y_2 W_3 + b_3 \tag{4}$$

where $W_3 \in \mathbb{R}^{h_2\times 3}$ and $b_2 \in \mathbb{R}^{1\times 3}$. As our goal is to do a 3-class classification task, Y3 represents the predicted classes of the corresponding inputs in a one-hot encoded form.

III. Results and Discussions

A. Experimental Setups

Using the idea of $N - 2$ security criteria, we simulate the cascading failure scenarios in the UIUC 150-Bus test system through setting any two different power lines as initial failed components which are to be removed from the test power network to construct our dataset. So, we obtain 20503 failure propagation profiles for training and testing the neural network-based classifier. For each of the failure propagation, the power matrix difference ΔM_p is calculated by setting $k = 2$ in (1), and the detection of the onset time is performed by adopting the method given in Liu *et al.* [9], i.e., $\Delta F_{\text{large}} = 10$ and $\Delta t_{\text{short}} = 10$ min. To fix three time intervals for three considered classes of the onset time, t_{c1} and t_{c2} are set to 10^2 min and 10^3 min from a practical perspective. Thus, 3704 non-urgent cases, 14700 urgent cases, and 2099 relative urgent cases can be found in 20503 data samples.

TABLE I: Average accuracy of training set, testing set, and two new scenarios including changing the original power parameters (Modified N-2 for short) and increasing k to 3 (N-3 for short), respectively.

Sets	Training	Testing	Modified N-2	N-3
Average accuracy	0.9417	0.8993	0.6332	0.8768

Based on the definition of our feature X, the dimension d is equal to 353 as there are 150 nodes and 203 transmission lines in the UIUC 150-Bus test system. For other settings, we set the number of hidden units $\{h_1, h_2\} = \{100, 50\}$, and the maximal epoch is 2000. Moreover, we use the classical stochastic gradient descent and cross-entropy loss to optimize our model, together with a weighted operation on the loss function through the reciprocal of the numbers of each class to reduce the influence of class imbalance. The ratio between the training set and the testing set are 0.7 and 0.3, respectively. All results are averaged from 5 different train-test splits.

B. Results

We first observe the change of accuracy of the training set and testing set with the increase of training epoch, as shown in Fig. 3. Here, the prediction accuracy is given by the ratio of the number of correctly predicted samples to the total number of predicted samples. We see that the training accuracy has a rapid rise at the very beginning, then steadily increases to around 90%. Moreover, the testing accuracy follows a similar trend as the training accuracy and reaches a high accuracy eventually, indicating that our model can effectively avoid the overfitting problem. From the second and third columns of Table I, we can conclude that the proposed method can achieve extremely good performance on both the training and testing sets.

Then, we consider two new challenging scenarios. From (1), we know that our model extracts the difference of power parameters between the original parameters in the initial state of the test power network and the latest parameters after removing k transmission lines. Therefore, we wish to assess the generality of the proposed method when we modify the original parameters $M_{p,N}$ (i.e., denoted as Modified $N-k$), and increase the number of k (i.e., denoted as $N - k$). To achieve this, we generate new samples and obtain the corresponding power difference matrix ΔM_p for each scenario. In particular, the additional testing data samples include two sets of 5000 scenarios considering $N - 2$ security criteria when the initial power demanded in loading nodes and power supplied by generating nodes are varied and $N - 3$ security criteria, respectively. It is worth noting that we directly conduct the test on the model that is trained with the original training set, instead of retraining a new model to fit the new generated samples.

According to the last two columns of Table I, our method performs well when the size of the attack area increases from k to 3, implying that our method can be generalized

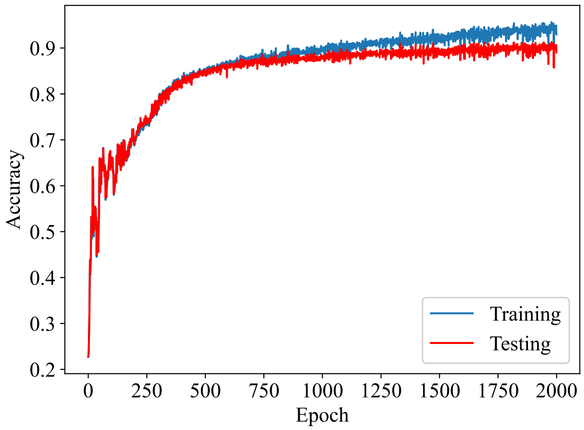

Fig. 3: Accuracy of training set and testing set with the increase of epoch.

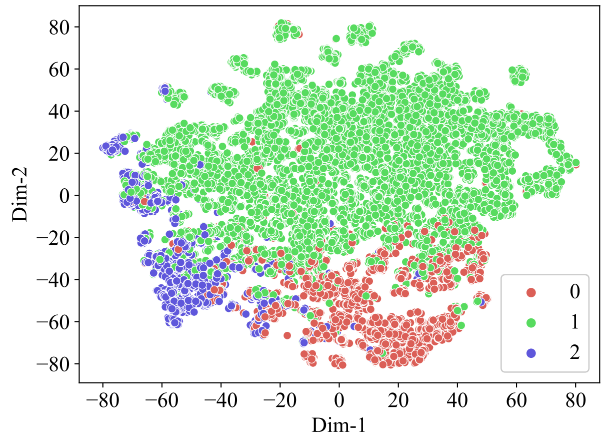

Fig. 4: t-SNE plot of the hidden vectors of each sample in 2D space. Representation of the samples from the same class are plotted with the same color. Here classes 0, 1, and 2 represent the non-urgent, urgent, and relatively urgent cases, respectively.

to different sizes of the attack area. However, the proposed method can only obtain an average of 63.32% accuracy in the scenario when the original power parameters $M_{p,N}$ are changed, which is a significant drop compared to the cases when using the original $M_{p,N}$. We believe a possible reason is that insufficient features are extracted based on the difference between the original power parameters and the latest power parameters after removing k lines. In our training model, all training samples share the same original power matrix $M_{p,N}$. However, as we modify the original $M_{p,N}$ in this new scenario, there might exist samples that have similar features ΔM_p but are obtained from different combinations of power parameter matrices $M_{p,N}$ and $M_{p,N-k}$, causing the confusion to our pre-trained model's decision. One possible solution in future work is to induce new features that are able to embrace the variation of the original power matrix $M_{p,N}$.

Finally, to illustrate the learning ability of the proposed method, we also examine the learning representation of each sample in the 2D space. Fig. 4 offers a visualization of the output of the second hidden layers (i.e., Y_2) on the 2D space by using t-SNE [18] technique. We can observe relatively clear boundaries between the samples with different classes, and the samples belonging to the same class usually cluster together. All of our experiments show that the proposed method can learn useful information from the difference of power information in the nodes and transmission lines, which contributes to the ability to infer the onset time as accurately as possible.

IV. Conclusions

An accurate prediction of the onset time of cascading failure in power systems on the one hand tells operators when a large amount of failure would occur within a short time, and on the other hand allows efficient assessment of the robustness of power systems against cascading failure. This paper demonstrates a neural network-based classifier that classifies the onset time into three separate time intervals, corresponding to urgent, relatively urgent, and non-urgent cases with around 90% accuracy by using only the initial state of the power network. With the promising ability of neural networks in learning the features given by the topological and electrical information of a power network, the trained classifier can predict the onset time in new testing datasets with acceptable accuracy. Future research will improve the accuracy of the neural network-based classifier with the aim to effectively replace the time-consuming simulation-based methods for assessing the robustness of large-scale power systems, especially when various operational states have to be considered. Besides the onset time, the sequence of failed components during failure propagation can also be predicted by using machine learning-based algorithms, which further helps prevent catastrophic power outages.

Acknowledgment: This work was funded by Hong Kong Research Grants Council under GRF 152150/19E, and City University of Hong Kong Grant 9380114.

References

[1] U.S.-Canada System Outage Task Force, "Final report on the August 14th blackout in the United States and Canada," US Dept. Energy and National Res. Canada, Tech. Rep., 2004.

[2] "Report on the Grid Disturbance on 30th July 2012 and Grid Disturbance on 31st July 2012," Enquiry Committee Constituted by Ministry of Power, Govt. of India, Tech. Rep., 2012.

[3] "Technical report on the events of 9 August 2019," 2019. [Online]. Available: https://www.ofgem.gov.uk/sites/default/files/docs/2019/09/eso_technical_report_-_final.pdf

[4] D. Liu, X. Zhang, and C. K. Tse, "A tutorial on modeling and analysis of cascading failure in future power grids," *IEEE Trans. Circ. Syst. II, Expr. Briefs*, vol. 68, no. 1, pp. 49–55, Jan. 2021.

[5] C. C. Chu and H. H. C. Iu, "Complex networks theory for modern smart grid applications: A survey," *IEEE J. Emerg. Sel. Top. Circ. Syst.*, vol. 7, no. 2, pp. 177–191, Jun. 2017.

[6] J. Fang, J. Wu, Z. Zheng, and C. K. Tse, "Revealing structural and functional vulnerability of power grids to cascading failures," *IEEE J. Emerg. Sel. Top. Circ. Syst.*, vol. 11, no. 1, pp. 133–143, Mar. 2020.

[7] J. Yang, C. K. Tse, M. Huang, and X. Fu, "Homoclinic bifurcation of a grid-forming voltage source converter," *IEEE Trans. Power Electron.*, vol. 36, no. 11, pp. 13 176–13 187, Nov. 2021.

[8] C. Kim, K. Kim, P. Balaprakash, and M. Anitescu, “Graph convolutional neural networks for optimal load shedding under line contingency,” in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2019, pp. 1–5.

[9] D. Liu, C. K. Tse, and X. Zhang, “Robustness assessment and enhancement of power grids from a complex networks perspective using decision trees,” *IEEE Trans. Circ. Syst. II, Expr. Briefs*, vol. 66, no. 5, pp. 833–837, Apr. 2019.

[10] Z. Zhang, R. Yao, S. Huang, Y. Chen, S. Mei, and K. Sun, “An online search method for representative risky fault chains based on reinforcement learning and knowledge transfer,” *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 1856–1867, May 2020.

[11] M. Abedi, M. R. Aghamohammadi, and M. T. Ameli, “Svm based intelligent predictor for identifying critical lines with potential for cascading failures using pre-outage operating data,” *Int. J. Electr. Power Energy Syst.*, vol. 136, p. 107608, Mar. 2022.

[12] R. R. Hossain, Q. Huang, and R. Huang, “Graph convolutional network-based topology embedded deep reinforcement learning for voltage stability control,” *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4848–4851, Sep. 2021.

[13] Z. Yan and Y. Xu, “A multi-agent deep reinforcement learning method for cooperative load frequency control of a multi-area power system,” *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4599–4608, Nov. 2020.

[14] X. Zhang, C. Zhan, and C. K. Tse, “Modeling the dynamics of cascading failures in power systems,” *IEEE J. Emerg. Sel. Top. Circ. Syst.*, vol. 6, no. 2, pp. 228–238, Jun. 2017.

[15] A. M. Farid, “LIINES smart power grid test case repository,” http://amfarid.scripts.mit.edu/Datasets/SPG-Data/index.php.

[16] F. Rosenblatt, “Principles of neurodynamics. perceptrons and the theory of brain mechanisms,” Cornell Aeronautical Lab Inc Buffalo NY, Tech. Rep., 1961.

[17] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning internal representations by error propagation,” California Univ San Diego La Jolla Inst for Cognitive Science, Tech. Rep., 1985.

[18] L. Van der Maaten and G. Hinton, “Visualizing data using t-SNE,” *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, Nov. 2008.