ScuDo

Scuola di Dottorato ~ Doctoral School

WHAT YOU ARE,TAKES YOU FAR

XXXII Cycle

Data-driven Analysis of Outages in Distribution Network

Yang Zhang

Supervisor: Prof. Ettore Bompard

Research context and motivation

- Unlike the power transmission system, the distribution network usually faces the challenges due to the large number of equipment, complex structure and various customers, which raise the possibilities of power supply interruptions to a large extent. In the meanwhile, the complicated causes and unclear relations bring difficulties to an accurate outage prediction. Since the consumers' satisfaction is deeply dependent on the number of interruptions, the Distribution System Operators (DSOs) have been seeking a reliable way to forecast the number of outages for a better maintenance plan and operations.
- With the fast development of digital technology and cloud computing, more and more data are produced through digital equipment and sensors, such as smart phones, computers, advanced measuring infrastructures, etc, as well as through human activities and communications.

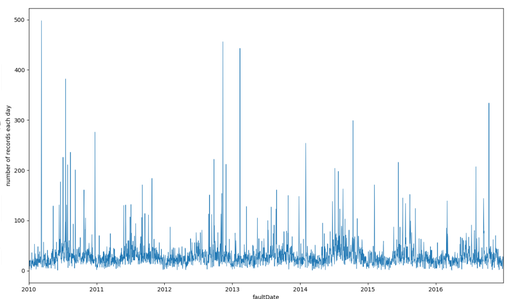

Figure 1 Number of interruption records in distribution network

- The great progress of information and communication technology (ICT) provides a new vision for engineers to perceive and control the traditional electrical system and makes it smart. An embedded information layer into the energy network produces huge volume of data, including measurements and control instructions in the grid for collection, transmission, storage and analysis in a fast and comprehensive way. It also brings a lot of opportunities and challenges to the data analysis platform.

Addressed research questions/problems

- There are some researches trying to analyze the impact of external factors to the outage in power grid in statistical models. In theory, with the information of all the potential causes collected, it becomes possible to build a regression model to predict the number of outages under a certain condition. For example, the adverse weather conditions are usually taken as an important factor for the reliability assessment in power system.
- However, due to the fact that most of the utilities focus more on the fast repair of electrical equipment rather than detailed investigation of failure causes, there are typically limited information to investigate the real reasons behind an outage. Moreover, as the majority of feeders in an urban distribution network, the underground cables are not as sensitive as overhead lines to the external factors, which weakens the performance of traditional regression models.

Novel contributions

- The grey model, which is a classic theory for uncertain issues with limited sample size and poor data information, is a promising solution for the outage prediction problems. As a frequently used prediction technique, GM(1,1) has been applied to various engineering problems, including the electrical load forecasting, wind power prediction, and so on. In our study, an improved grey model was adopted to forecast the number of outages due to the absence of causes and external information.
- Compared to the typical basic grey model, the GA optimization was introduced to determine the proper parameter with an objective as the minimum error between predictions and real values. Furthermore, a rolling mechanism was applied to track the latest trend of data dynamically and improve the precision of estimation.

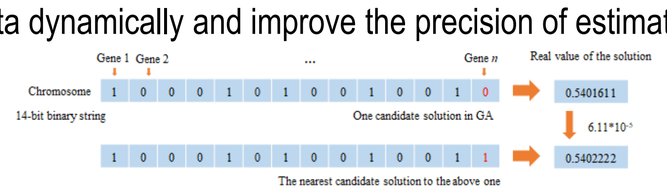

Figure 2 An example of decoding process in GA

Submitted and published works

- Yang Zhang, Andrea Mazza, Ettore Bompard, et al, "Prediction of Power Outages in Distribution Network with Grey Theory", IEEE Conference SEST, Porto, September 9-11, 2019
- Yang Zhang, Andrea Mazza, Ettore Bompard, et al, "Data-driven Feature Description of Heat Wave Effect on Distribution System", IEEE Conference on PowerTech, Milano, June 23-27, 2019
- Yang Zhang, Colella Pietro, Andrea Mazza, et al, "Abnormal Electricity Consumption Detection from Incomplete Records in Power System", IEEE Conference ISGT Asia, Chengdu, May 21-24, 2019

Adopted methodologies

- Basic GM(1,1) Model
- Grey model is a classic method for studying the trend from discrete data series with limited samples and inadequate information. By accumulating the original data series, the randomness between samples could be reduced and a clear trend is possible to be revealed. Therefore, the first step of grey model is to add up the values:

$$\mathbb{X}^{(1)}(k) = \sum_{i=1}^{k} \mathbb{X}^{(0)}(i) \qquad k=1, \ldots, N$$

- The k-th background value $\mathbb{Z}(k)$ is defined as the weighted average value of the k-th and (k-1)-th accumulated data as below

$$\mathbb{Z}(k) = \mu\mathbb{X}^{(1)}(k-1) + (1-\mu)\mathbb{X}^{(1)}(k)$$

- The first-order grey differential equational of variable x is given as follows:

$$\mathbb{X}^{(0)}(k) + a\mathbb{Z}(k) = b \qquad k=2, \ldots, N.$$

- This equation corresponds to the white differential equation below based on the assumption that the accumulated sequence increase exponentially.

$$\frac{\mathrm{d}\mathbb{X}^{(1)}(t)}{\mathrm{dt}} + a\mathbb{Z}(t) = b$$

- where a and b are the development coefficient of the system and endogenous control grey scale, respectively.
- By rewriting the above equation in the matrix format, we could get the following

$$\begin{bmatrix} \mathbb{X}^{(0)}(2) \\ \mathbb{X}^{(0)}(3) \\ \vdots \\ \mathbb{X}^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -\mathbb{Z}(2) & 1 \\ -\mathbb{Z}(3) & 1 \\ \vdots & \vdots \\ -\mathbb{Z}(n) & 1 \end{bmatrix} \times \begin{bmatrix} a \\ b \end{bmatrix}$$

- The parameters a and b could be estimated with the least square method. Then the direct output of the grey prediction model is the estimation of the accumulated values:

$$\hat{\mathbb{X}}^{(1)}(k+1) = \left(\mathbb{X}^{(1)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a}$$

- The prediction of the original data could be determined

$$\hat{\mathbb{X}}^{(0)}(k+1) = \hat{\mathbb{X}}^{(1)}(k+1) - \hat{\mathbb{X}}^{(1)}(k)$$

- Optimization of μ with GA
- In traditional grey model GM(1,1), the weight μ for calculating background values are fixed as 0.5, which means that the background values are equal to the average of two continuous accumulated data. However, the value of μ is possible to be within the interval [0,1]. An optimization of μ could be carried out for building a better grey model. In this paper, the GA technique is used for finding the optimized value of μ and improve the flexibility of the grey model.
- There are two 14-bit chromosomes composed by binary genes shown in Figure 1. In our case, the value of μ is within the range [0,1], therefore the real values of the candidate solutions can be decoded by locating the positions represented by the chromosomes in [0,1]. The difference of the real values between the two nearest chromosomes refers to the precision of the final solution with GA. With 14-bit chromosomes, the minimum real value of the difference between two candidate solutions is $6.11 * 10^{-5}$, which indicates that the precision of the solution is 4 decimals.

Conclusion and future work

- In our study, the outage in the local urban distribution network have been collected and used for the building of grey model. From 2008 to 2012, the number of outages in every month is shown in Figure 3. As can be seen from the figure, the monthly outage number increases during summer and decrease in winter.
- With the objective function as the minimum errors between the model and real values, even limited number of records could be used to build the prediction model.
- In the future work, the potential impetus for the increasing of outages during the years needs to be analyzed with more data sources.

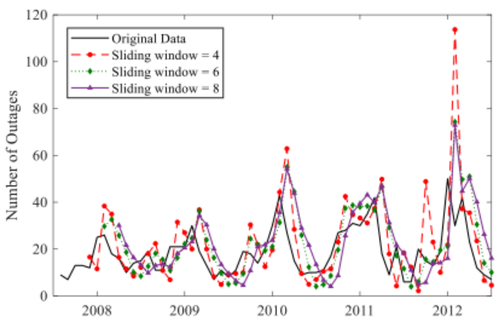

Figure 3 Results of monthly outage prediction with different sliding windows

POLITECNICO DI TORINO

PhD program in

Electrical, Electronics and Communications Engineering