Scientific
Research

Branching Process based Cascading Failure Probability Analysis for a Regional Power Grid in China with Utility Outage Data

Hui Ren[1], Ji Xiong[1], David Watts[2,3], Yibo Zhao[4]

[1]Department of Electrical Engineering, North China Electric Power University, Baoding, China
[2]Pontificia Universidad Catolica de Chile, PUC, Vicuna Mackena 4860, Macul, Santiago, Chile
[3]University of Wisconsin-Madison, Wisconsin, USA
[4]School of information & Electronic Engineering, Beijing Institute of Technology, Beijing
Email: hren@ncepubd.edu.cn, dwatts@ing.puc.cl

ABSTRACT

Studying the propagation of cascading failures through the transmission network is key to asses and mitigate the risk faced the energy system. As complex systems the power grid failure is often studied using some probability distributions. We apply 4 well-known probabilistic models, Poisson model, Power Law model, Generalized Poisson Branching process model and Borel-Tanner Branching process model, to a 14-year utility historical outage data from a regional power grid in China, computing probabilities of cascading line outages. For this data, the empirical distribution of the total number of line outages is well approximated by the initial line outages propagating according to a Borel-Tanner branching process. Also for this data, Power law model overestimates, while Generalized Possion branching process and Possion model underestimate, the probability of larger outages. Especially, the probability distribution generated by the Poisson model deviates heavily from the observed data, underestimating the probability of large events (total no. of outages over 5) by roughly a factor of 10-2 to 10-5. The observation is confirmed by a statistical test of model fitness. The results of this work indicate that further testing of Borel-Tanner branching process models of cascading failure is appropriate, and should be further discussed as it outperforms other more traditional models.

Keywords: Cascading Failure; Poisson; Power Law; Branching Process; Generalized Poisson; Borel-Tanner

1. Introduction

Cascading failure is the process by which initial outages of components of the electric power transmission system can occasionally propagate to more widespread outages and large blackouts. These blackouts involve complex chains of events, that are not completely independent, as an outage that have already occurred weakens the system, making the systems more vulnerable and further outages more likely [1].

Since 2001, researchers with multiple backgrounds of electrical engineering, mathematics, physics and nonlinear dynamics have work more heavily addressing this topic from different angles. This research on cascading models is surveyed and presented in [2,3], and is mainly focused on 1) studying on the evolution of cascading failure from a long term point of view; 2) studying on the effect of specific disturbances on cascading propagation; 3) identifying the vulnerable area of the system according to the topological features; 4) evaluating hazards the initial failure could cause to the system by identifying the afterwards cascades being initiated. All these researches provide helpful research results.

Dealing with these issues and performing risk assessment requires often identifying the probability distribution of blackout size. It has been widely observed that the probability distributions produced by models in above-mentioned research works have approximate power law regions that cannot be produced by independent outages and are broadly consistent with the characteristics observed in the utility historical outage data from several countries.

Historical outage data has always been used to power system reliability evaluation, considering it encompasses the effect of both inadequacy and insecurity of the studied system. A bulk statistical approach based on the historical outage data is different from traditional methods of risk analysis that rely on detailed analysis of enumerated interactions, and will be complementary to traditional methods, especially on the risk analysis of large blackouts because of the challenges on computational power faced by enumerated interaction methods [1] and

the difficulties addressing complexity and non-independency among different events.

1.1. Probabilistic Models for Utility Outages

While the literature body in this field is very broad, this section summarizes previous work on probabilistic models for utility outages that relates more closely to our work.

Chen and McCalley, after analyzing the North American transmission line outaged data in [4], proposed an accelerated propagation model for the number of transmission line outages [5]. They try a generalized Poisson distribution and a negative binomial distribution for the accelerated propagation model. The fitness test of the accelerated propagation model to the reference utility outage data showed, for the PDF of the number of line outaged, the proposed exponentially accelerated cascading model on the probability of outages with more than 6 lines fit better than other models (0.00061 for accelerated propagation model, while 0.00053 for the observed outage data, probabilities produced by Generalized Poisson model and Negative Binomial model are one order of magnitude smaller that that of the observed data).

In [1], Ren and Dobson demonstrated a way to test a branching process based bulk statistical model of cascading line outages on industry data. The model described the development of cascading failure starting from initial outages that then propagates in stages. The average amount of propagation was estimated and hence the probability distribution of the size of the cascading failure (measured by the total number of line outages) was predicted. The fit of the model is examined with a 9-year historical outage data set from a regional power grid. Furthermore, Dobson [6] improves the branching process model in [1] by considering variant propagation in different stages, modeling the increases of propagation as the cascade proceed following the observation in the utility data [8]. The branching process model is then used to predict the distribution of total number of outages for a given number of initial outages. They study how the total number of lines outaged depends on the propagation as the cascade proceeds.

1.2. Proposed Research

The power system is carefully designed and operated so that most transmission line outages do not propagate over the system and it only has to encompass one or few outages occurring together. This paper deals exactly with those situations where failure may propagate, protections, controls and other factors may not operate as desired, and several components of the system may be affected. As previous references, failure propagation is modeled using probabilistic models, allowing for multiples failures modes.

Power grids differ one from another mainly in topology, size, capacity, interconnectedness, loading level, and several other physical/technical features. The planning and operation of the market or the power system also varies, but not as widely as physical features (dispatch rules, unit commitments, market clearing processes differs but have similar goals and results).

Although these differences exist, outage data seems to suggest similar cascading failure behavior in widely different systems. As such, research is still undergoing aiming to identify the best probabilistic models for cascading modeling in power systems. In this paper, we compare the fitness of 4 different statistical models to the 14-year utility outage data from a regional power grid.

This paper is organized as follows. In section II, outage data sources and handling procedures to group failures are presented. In section III three probabilistic models ever used for transmission line outage analysis are introduced first. Utility outage data is introduced in section III, and also the simple grouping techniques of utility outage data for branching process based analysis.

2. Utility Outage Data

Cascading failure models are very constrained by outage data availability and some models require some pre-possessing or handling the data. This section address these topics, describing different data sources, the data set used here and the handling procedures required to fit the models.

2.1. Main Sources of Utility Outage Data

For the researches on cascading large blackouts, the utility outage data are valuable references on modeling and useful diagnostics in monitoring the process and extent of blackouts. Since data is scarce, here we summarize the data available from cited references.

Ref. [4] summarized the results of a survey of design characteristics of and outage experience with overhead transmission at voltage 230kV and above in USA and Canada, for a purpose of providing technical support on the probabilistic system models for planning and operation. The outage data were submitted by utilities from all nine NERC/USA reliability regions and by the Canadian Electric Association representing all of Canada. The statistics provided classifications and analysis on 38,489 outage data according to voltage level, region, causes, or duration time, etc.

Every year, the North American Electrical Reliability Council (NERC) publishes a documented list summarizing major disturbances [7], which is used as strong evidence in abovementioned researches for bigger probability of large blackouts.

Ref. [8] provides a historical outaged data recorded by a North American utility over a period of 12.4 years and is being upgraded constantly. The transmission line outage data is required by NERC for the Transmission Availability Data System (TADS). The data for each transmission line outage includes the outage time (to the nearest minute) as well as other data. All the line outages are automatic trips. More than 99% of the outages are of lines rated 69kV or above and 97% of the outages are of lines rated 115kV or above. There are several types of line outages in the data and a variety of reasons for the outages [6].

Our previous research and the work of this paper are based on 9-year historical outage data in [1] and 14-year outage data in this paper, which are extracted from the fault data information recorded by protective relays in a regional power grid. The data we can reach provides the voltage level and the time stamp of the failure, the sequence of the failures, and failure cause, however with on specific information on load shed. The data set is described in details in next section.

2.2. Utility Outage Data Used in this Paper

The outage data used in this paper is from a regional electric power transmission system with, approximately, 190 220kV buses, and 12 buses at 500kV. The data is recorded over about 14 years staring in 1997 and ending in 2011.

The historical outage data is extracted from the report containing the outage details recorded by fault recorder devices and the fault analysis by relay engineers. The extracted information includes the contingency type (transmission line failure, or busbar failure, or generator failure…), contingency time (to the nearest minute), voltage level, and the auto-recloser's action. In this research, only transmission line outages in 220kV and 500kV are computed in the model. Outages at lower voltage levels are not considered because of the potential number of unrecorded events.

When processing the outage data, the voltage level, line outage type (single phase or three-phase and detailed causes of the line outages are neglected, and are regarded as the same. The neglecting in the bulk statistical analysis is appropriate, for no matter what the details on causes, type or voltage levels are, and they all result in the weakness of the transmission system to various extent. The more sever the outage is (outages on higher voltage level, or more initial outages, or poor operation and control techniques), the more sever outcomes follows, such as even more outages, while bulk statistical analysis based on the historical outage data actually takes all of these into consideration, and lead to a more credible evaluation on system's risk level.

Large flashover events in the data with approximately 260 outages over two days are neglected because of lacking time tags. There is not large cascading blackout (totally blackout) in this regional power grid over the analyzed horizon. Data recording is not perfect and there is some kind of incomplete outage data recording, those should be smaller failures, not big ones. Therefore the tail part in the probability distribution (*PDF*) versus outages with various sizes should be credible.

The clustering of outages in stages can be seen in **Figure 1**. The x-axis shows time since start of cascade for outages in each of the 458 cascades. The first min of each cascade is shown. Multiple outages at the same time are shown slightly displaced.

2.3. Data Grouping by Stages

The processing in this section for the outage data is mainly for the application of branching process model. Here, we follow the data grouping technique in [1]. The outage data are grouped according to their time stamps without considering the geographical information of the outage. Although successive outages in areas far way from each other may not have direct electrical connections, but the successive happening does suggest a further weakening of the system. Moreover, when the data is in bulk and the system is in large-scale, the difficulty from considering more detail information of each outage would compromise the application of the method.

Successive outages separated in time by more than one hour are assumed belong to different cascades, for operator actions are usually completed within one hour. Successive outages in a given cascade separated in time by more than one minute are assumed in different stages within that cascade, for transients or auto-recloser actions are completed within one minute.

733 outages are abstracted from the failure records, and from them, 459 cascades are obtained by the introduced grouping technique. **Table 1** is obtained by summing over all the 459 cascades the number of outages in
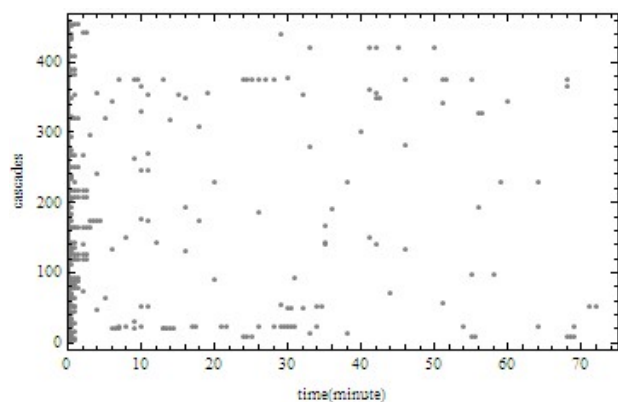

Figure 1. Clustering of outages in stages.

each stage. That is, of the 733 outages, 556 are in stage 0 of a cascade (for some cascades, there are more than 1 failures in stage 0), 83 are in stage 1 of a cascade, and so on. Failures in stage 0 are defined belong to initial failure of a cascade in the following analysis.

Table 2 and **Table 3** give the statistics of the no. of lines lost in each initial failure, and in each cascade, respectively. *N-r* in **Tables 2** and **3** means the loss of r transmission lines in the power system (a traditional reliability event indicator). The number of line outages in initial failures could be as large as 7, and the total number of outage in the cascade is at most 19, however the probability of these severe occasions are extremely small.

Table 1. Number of outages in each stage summed over the cascades.

Stage	Failure No.	Stage	Failure No.	Stage	Failure No.	Stage	Failure No.
Z_0	556	Z_4	14	Z_8	3	Z_{12}	2
Z_1	83	Z_5	6	Z_9	3	Z_{13}	1
Z_2	31	Z_6	5	Z_{10}	3	Z_{14}	1
Z_3	20	Z_7	3	Z_{11}	2	Z_{15}	0

Table 2. Statistics on the no. of transmission line lost in the initial failure.

Con. Type of initial failure	No.	Probability
N-1	402	0.8758
N-2	35	0.07625
N-3	13	0.02832
N-4	5	0.01089
N-5	0	0
N-6	3	0.006536
N-7	1	0.002179

Table 3. Statistics on the total no. of transmission line lost in a cascade.

Con. Type of the cascades	Total no.	Probability
N-1	341	0.7429
N-2	62	0.1351
N-3	27	0.05882
N-4	10	0.02179
N-5	4	0.008715
N-6	5	0.01089
N-7	5	0.01089
N-8	1	0.002179
N-9	1	0.002179
N-10	1	0.002179
N-16	1	0.002179
N-19	1	0.002179

The probability distribution (*PDF*) of initial and total line outages are shown in **Figure 2** on a log-log scale. Raw data is shown in **Figure 2** with no binning. It has a peak at 6 and 7 outages. One reason for this is that some cascades are initiated by a bus outage, and the relay trips off all transmission lines connected to that bus simultaneously at the start of the cascade. On log-log scale, the PDF is roughly a straight line, showing a "heavy tail", consistent with researches on cascading blackouts in [10].

3. Assessing Four Probabilistic Models

We introduce in this section four different probability models which we use to estimate the distribution of the total number of line outages in this paper. Among these models, branching process model is introduced as a focal point and will be compared with other traditional probability models using fitness test.

3.1. Traditional Poisson Model [9]

Consider a random variable $T \in \{0,1\}$, with $T = 1$ representing the event of an individual line tripping and the probability is $P(T = 1) = p$. Then the probability of tripping of each line can be represented as

$$P(T = t \mid p) = p^t (1-p)^{1-t}, t = 0,1; 0 \leq p \leq 1 \quad (1)$$

Suppose that the total number of lines in a power system is *N*, and each line has the same probability *p* to be tripped and each trip event is independent of any other one. Then the probability distribution of total number of line outages Z subjects to the binomial distribution:

$$P[Z = r] = C_N^r p^r (1-p)^{N-r}, r = 0,1,2,\cdots,N \quad (2)$$

Consider that *N* is large and *p* is small under general conditions, then the formula can be approximated by the Poisson distribution with a parameter of $\lambda_{con} = Np$:

$$P[Z = r] = C_N^r p^r (1-p)^{N-r}$$
$$\approx e^{-\lambda_{con}} \lambda_{con}^r / r!, r = 0,1,2,\cdots,N \quad (3)$$
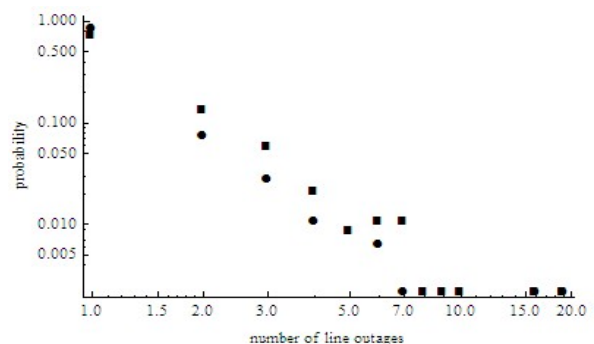

Figure 2. Probability distribution of initial (dots) and total (squares) line outages.

In this paper, we use maximum likelihood estimation [15] to estimate the parameter λ_{con}. If the space X is discrete with probability distribution function

$$P(X = x) = p(x; \theta),$$

then the joint probability distribution function of event $\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\}$ can be shown as

$$L(\theta) = L(x_1, \cdots, x_n; \theta) = \prod_{i=1}^{n} p(x_i; \theta) \quad (4)$$

where θ is the model parameter to be estimated. The parameter value $\hat{\theta}$ which makes (4) reach the maximum value is called as maximum likelihood estimator, we can find out $\hat{\theta}$ by mathematical methods such as derivation.

We can get $\hat{\lambda}_{con} \approx 0.5970$ based on **Table 2** and maximum likelihood estimation. The estimated distribution is as follows:

$$P[Z = r] = e^{-0.5970} 0.5970^{r-1} / (r-1)!, r = 1, 2, \cdots \quad (5)$$

3.2. Power Law Model [10-11]

The total number of line outages Z has a normalized distribution as shown in (6) when it follows a power law:

$$P[Z = r \mid q] = r^{-q} / \sum r^{-q}, q > 0; r = 1, 2, \cdots \quad (6)$$

When we draw the relationship of $P[Z = r]$ and r on a log-log plot, we will find a straight line with slope $-q / \sum r^{-q}$, this indicates the distribution follows a power law. Using maximum likelihood estimation introduced above, we can get $\hat{q} \approx 2.0$ based on **Table 2**, then the expression of using power law model to estimate the distribution of outage lines can be shown as follows:

$$P[Z = r] = r^{-2.0} / \sum_{k=1}^{\infty} r^{-2.0} = 0.645 r^{-2.0}, r = 1, 2, \cdots \quad (7)$$

3.3. Branching Process Model [1]

The propagation λ which means the mean number of child failures for each parent failure plays a critical role in branching process. We consider λ as an invariant constant in this paper, the positive number of failures in stage zero Z_0 will produce outages in the next stage with a mean number of λZ_0, the child failures then become parents to produce outages with a mean number of $\lambda^2 Z_0$ in the next stage and so on [12]. Here we use the method introduced in [1] to estimate λ:

$$\hat{\lambda} = \frac{\sum_{i=1}^{J} (Z_1^{(i)} + Z_2^{(i)} + \cdots + Z_{N(i)}^{(i)})}{\sum_{i=1}^{J} (Z_0^{(i)} + Z_1^{(i)} + \cdots + Z_{N(i)-1}^{(i)})} \quad (8)$$

where J represents the total number of line outages,

$N(i)$ represents the maximum stage with nonzero failures. Based on **Table 1** and (8), we can know

$$\hat{\lambda} = \frac{Z_1 + Z_2 + \cdots + Z_{14}}{Z_0 + Z_1 + \cdots + Z_{13}} \approx 0.24 \quad (9)$$

3.2.1. Generalized Poisson model

If the distribution of initial failure numbers follows a Poisson distribution with the parameter θ and ignores the condition with zero initial failure, the distribution of initial numbers is as follows:

$$P[Z_0 = r] = \frac{e^{-\theta} \theta^r}{(1 - e^{-\theta}) r!}, r = 1, 2, \cdots \quad (10)$$

The mean number of initial failures can be shown

$$\overline{Z_0} = \hat{\theta} / (1 - e^{-\hat{\theta}}) \quad (11)$$

We can get $\overline{Z_0} \approx 1.211$ based on **Table 2**, and then we get the estimator $\hat{\theta} \approx 0.3965$ using (11). The distribution of the total number of line outages under this condition follows generalized Poisson distribution [13] as shown in (12).

$$P[Z = r] = \theta (r\lambda + \theta)^{r-1} \frac{e^{-r\lambda - \theta}}{(1 - e^{-\theta}) r!} =$$

$$0.3965 (0.24r + 0.3965)^{r-1} \frac{e^{-0.24r - 0.3965}}{(1 - e^{-0.3965}) r!}, \quad (12)$$

$$r = 1, 2 \cdots$$

3.2.2. Borel-Tanner model

If the initial failure numbers follows arbitrary distribution, then the total number of line outages follows a Borel-Tanner distribution [14].

$$P[Z = r] = \Sigma_{z_0=1}^{r} P[Z_0 = z_0] z_0 \lambda (r\lambda)^{r-z_0-1} \frac{e^{-r\lambda}}{(r - z_0)!}$$

$$= \Sigma_{z_0=1}^{r} 0.24 P[Z_0 = z_0] z_0 (0.24r)^{r-z_0-1} \frac{e^{-0.24r}}{(r - z_0)!} \quad (13)$$

$$r = 1, 2, \cdots; z_0 = 1, 2, \cdots r$$

where $P[Z_0 = z_0]$ can be calculated based on the data given in **Table 2**.

4. Comparison Analysis of 4 Probabilistic Models

We evaluate the expressions in (5), (7), (12), and (13) for the loss of different number of lines, i.e.

$$r = \{1, 2, 3, 4, 5, \ldots, 10\}.$$

Probabilities obtained from 4 models and recorded outage data are listed in **Table 4**, where G-P represents Generalized Poisson, and B-T represents Borel-Tanner. The

probabilities in **Table 4** are given in **Figrue 3**. **Figrue 3** shows Log-Log plots of PDFs in Poission model in (5) (Inverse triangles), Power Law model in (7) (triangles), Generalized Poisson branching process in (12) (squares), Borel-Tanner branching process in (13) (diamonds) and observed data (dots).

For outages fewer than 5, all models fit well (as shown in **Figure 3**). However, Borel-Tanner branching process model is better than the other 3 models for outages bigger than losing 5 lines. Power law overestimates the probability of larger outages. The curve generated by the Poisson model deviates heavily from the observed data, underestimating the probability of large events (r>5) by roughly a factor of 10^{-2} to 10^{-5}.

In order to get a more clear view, the probability distribution of total number of outages from observed data and estimated using Borel-Tanner branching process are given in **Figure 4**. We can see that the distribution estimated using Borel-Tanner branching process can fit the distribution from data well, and in the log-log plot it also shows a heavy tail.

Table 4. Probabilities from 4 Probabilistic Models and Observed Data.

| r	No. of failure	Observed data	Poisson	Power law	Branching process G-P	Branching process B-T
1	341	0.7429	0.5505	0.6450	0.6398	0.6877
2	62	0.1351	0.3286	0.1613	0.2211	0.1776
3	27	0.05882	0.09809	0.07167	0.08276	0.06875
4	10	0.02179	0.01952	0.04031	0.03273	0.03128
5	4	0.008715	0.002913	0.02580	0.01347	0.01419
6	5	0.01089	0.0003479	0.01792	0.005707	0.007945
7	5	0.01089	0.00003461	0.01316	0.002474	0.005052
8	1	0.002179	2.952E-06	0.01008	0.001092	0.003183
9	1	0.002179	2.203E-07	0.007963	.0004890	0.001912
10	1	0.002179	1.461E-08	0.006450	.0002217	0.001098
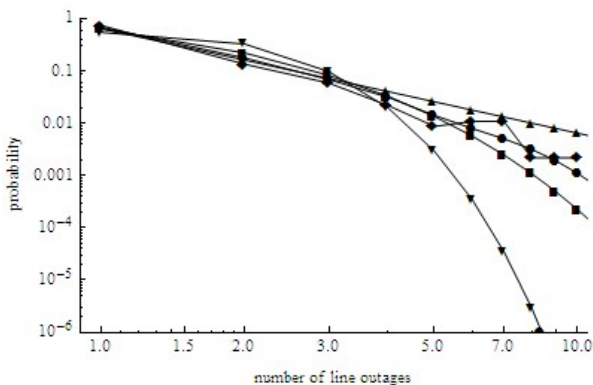
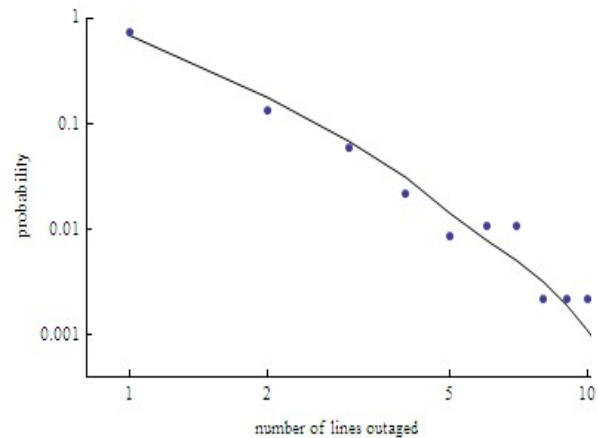**Figure 3. Log-Log plots of PDFs of 4 probabilistic models.**

Figure 4. PDFs of total No. of outages from observed data (dots) and by Borel-Tanner branching process (line).

Chi-square test is then used to provide a quantitative comparison among 4 models, and the test result is given in Table V. The Chi-square test [21] is widely used in statistics to test the fitness of a probability model to sample data. Suppose X follows the discrete distribution with the possible values of $1, 2, \cdots, k$, and the probability can be shown as $P(X = i) = p_i, 1 \le i \le k$. If we perform n trials, and the event $X = i$ results $X_i (i = 1, 2, \cdots, k)$ times, then we can use the statistic χ^2 to show how much the samples deviate from the distribution to be tested:

$$\chi^2 = \sum_{i=1}^{k} \frac{(X_i - np_i)^2}{np_i} \quad (14)$$

when $n \to \infty$, χ^2 follows the chi-square distribution χ^2 with the freedom number $k - 1$. We can obtain from (14) that the larger the statistic χ^2, the larger the deviation. Since the condition to let the conclusion introduced above be true is that the distribution must be a polynomial distribution and all of the np_i should be larger than 5, we decompose the sample space into 5 exclusive sets

$$M_1 = \{1\}, M_2 = \{2\}, M_3 = \{3\},$$
$$M_4 = \{4\}, M_5 = \{5, 6, \cdots\}.$$

The test result is shown in **Table 5**. If there are s estimated parameters in the probability model, the freedom number of the χ^2 distribution should be changed to $k - s - 1$. For instance, for the Borel-Tanner model, λ is a parameter to be estimated, so the freedom number of chi-square distribution is 5-1-1=3; but for the generalized Poisson model, the parameters that should be estimated are λ and θ, so the freedom number is changed to be 5-2-1=2. By inspecting **Table 4**, the value of χ^2 calculated from the Borel-Tanner model is obviously smaller than the values gotten from other models, that illustrates the Borel-Tanner model is far more fit than the other

Table 5. χ^2 Test Results For all models.

r	X_i	B-T		G-P		Power Law		Poisson	
		p_i	np_i	p_i	np_i	p_i	np_i	p_i	np_i
1	341	0.6877	315.7	0.6398	293.7	0.6450	296.1	0.5505	252.7
2	62	0.1776	81.5	0.2211	101.5	0.1613	74.04	0.3286	150.8
3	27	0.06875	31.56	0.08276	37.99	0.07167	32.90	0.09809	45.02
4	10	0.03128	14.36	0.03273	15.02	0.04031	18.50	0.01952	8.960
>= 5	19	0.03467	15.91	0.02361	10.84	0.08172	37.51	0.00329	1.510
m-r-1		3		2		3		3	
χ^2		9.31		34.07		22.87		292.4	

models. When we set the significance level be 0.01, the χ^2 of the Borel-Tanner model satisfies

$$\chi^2 < \chi^2_{0.01}(3) = 11.34 ,$$

that illustrates the Borel- Tanner model can exactly describe the distribution of real data under a significance level of 0.01.

5. Conclusions

We apply 4 probabilistic models, Poisson model, Power Law model, Generalized Poisson Branching process model and Borel-Tanner Branching process model, to a 14-year historical outage data from a regional power grid in China, computing probabilities of cascading line outages. We group the line outages into cascades and stages according to their outage times. For this data, the empirical distribution of the total number of line outages is well approximated by the initial line outages propagating λ according to a Borel-Tanner branching process with propagation parameter. For this data, Power law model overestimates the probability of larger outages, while the probability distribution generated by the Poisson model deviates heavily from the observed data, underestimating the probability of large events (outage no. greater than 5) by roughly a factor of 10^{-2} to 10^{-5}. The observation is confirmed by a statistical test of model fitness. The results of this work justify further testing of Borel-Tanner branching process models, leading to a promising research avenue for cascading failure modeling.

6. Acknowledgements

REFERENCES

[1]   H. Ren and I. Dobson, "Using Transmission Line Outage Data to Estimate Cascading Failure Propagation in an Electric Power System," *IEEE Transactions on Circuits and Systems*, Vol. 55, No. 9, 2008, pp. 927-931. doi:10.1109/TCSII.2008.924365

[2]   IEEE PES CAMS Task Force on understanding, prediction, mitigation and restoration of cascading failures, Initial review of methods for cascading failure analysis in electric power transmission systems," *IEEE Power Engineering Soc. General Meeting*, Pittsburgh PA USA, Jun. 2008.

[3]   D. Watts and H. Ren, "Classification and Discussion on Methods for Cascading Failure Analysis in Transmission System," ICSET, Singapore, Nov. 2008.

[4]   R. Adler, S. Daniel, C. Heising, M. Lauby, R. Ludorf and T. White, "An IEEE Survey of US and Canadian Overhead Transmission Outages at 230 kV and above," *IEEE Transactions on Power Delivery*, Vol. 9, No. 1, 1994, pp. 21 -39.doi:10.1109/61.277677

[5]   Q. Chen, C. Jiang, W. Qiu and J. D. McCalley, "Probability Models for Estimating the Probabilities of Cascading Outages in High-voltage Transmission Network," *IEEE Transactions on Power Systems*, Vol. 21, No.3, August 2006, pp.1423-1431. doi:10.1109/TPWRS.2006.879249

[6]   I. Dobson, "Estimating the Propagation and Extent of Cascading Line Outages from Utility Data with a Branching Process," *IEEE Transactions on Power Systems*, Vol. 27, No.4, 2012, pp. 2146-2155. doi:10.1109/TPWRS.2012.2190112

[7]   Information on Electric System Disturbances in North America can be Downloaded from the NERC Website at http://www.nerc.com/dawg/database.html.

[8]   Bonneville Power Administration Transmission Services Operations & Reliability website http://transmission.bpa.gov/Business/Operations/Outages.

[9]   Q. Chen and J. D. McCalley, "A Cluster Distribution as a Model for Estimating High-order Event Probabilities in Power Systems," 8th International Conference on Probabilistic Methods Applied to Power Systems. Ames, Iowa, 2004, 622-628.

[10] B. A. Carreras, D. E. Newman, I. Dobson and A. Poole, "Initial Evidence for Self-organized Criticality in Electric Power System Blackouts," *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences,* 2000, pp. 1411-1416.

doi:10.1109/HICSS.2000.926768

[11] I. Dobson, J. Chen, J. Thorp, B. A. Carreras and D. E. Newman, "Examining Criticality of Blackouts in Power System Models with Cascading Events," *Proceedings of the 35th Annual Hawaii International Conference on System Sciences,* 2002, pp. 803-812. doi:10.1109/HICSS.2002.993975

[12] I. Dobson, B. A. Carreras and D. E. Newman, "A Branching Process Approximation to Cascading Load-dependent System Failure," The 37th HICSS. 2004.

[13] I. Dobson, K. R. Wierzbicki, B. A. Carreras, V. E. Lynch and D. E. Newman, "An Estimator of Propagation of Cascading Failure," the 39th HICSS．2006．

[14] I. Dobson, K. Wierzbicki, B. A. Carreras, V. Lynch and D. E. Newman, "An Estimator of Propagation of Cascading Failure," The 39th HICSS, 2006.

[15] R. E. Walpole, Probability & Statistics for Engineers & Scientists, Seventh Edition. Pearson Education Asia Limited and Tsinghua University Press.