

基于历史故障记录数据的电网连锁故障规模概率分布研究

任惠¹, 熊吉¹, David Watts², 陈曦³

(1. 华北电力大学电气与电子工程学院, 河北 保定 071003; 2. Pontificia Universidad Catolica de Chile, PUC, Vicuna Mackena 4860, Macul, Santiago, Chile; 3. 西安理工大学水电学院, 陕西 西安 710048)

摘要: 连锁故障规模的概率分布描述了电网连锁故障的传播特点, 是衡量电网发生大规模停电故障概率的有效方法之一。针对历史故障统计数据计算, 是传统电力系统可靠性评估方法之一。将其与分支过程模型结合, 用于区域电网的连锁故障分析。采用某区域电网14年历史故障数据为样本数据, 针对多种概率模型进行比较分析, 提出采用波雷-坦尔分支过程模型计算该区域电网连锁故障规模的概率分布, 并采用误差分析研究了波雷-坦尔模型应用于实际电网风险管理的有效性和可能性。结果表明, 波雷-坦尔模型能够很好地估计线路故障规模的概率分布。在相同置信度要求下, 基于波雷-坦尔模型估计故障概率分布所需样本数据比直接根据实际故障数据计算所得概率分布所需样本数据降低一个数量级。

关键词: 连锁故障; 分支过程; 波雷-坦尔; 误差分析; 概率分布; 历史故障记录数据

Study on the historical outage data based probability distribution of cascading line failures in a regional power grid

REN Hui¹, XIONG Ji¹, David Watts², CHEN Xi³

(1. School of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China; 2. Pontificia Universidad Catolica de Chile, PUC, Vicuna Mackena 4860, Macul, Santiago, Chile; 3. College of Water and Electricity Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: The probabilistic distribution of cascading outages is one of the main measures to describe the propagation of cascading outages, and to evaluate the risk of the large scale outages of the power system. Historical outage data has always been used for power system reliability evaluation, and by combined with the branching process model, it is used for cascading outage analysis for a regional power grid. Based on the 14-year utility historical outage data from a regional power grid in China, several known probabilistic models are tested and compared, and a Borel-Tanner branching process model is proposed to estimate the probabilities of cascading line outages. Statistical error analysis is performed to study the effectiveness of applying the Borel-Tanner model to practical grid risk management. Results indicate that the empirical distribution of the total number of line outages is approximated well by the Borel-Tanner model. For the same confidential level, the estimation of the probability distribution of the larger cascades by the Borel-Tanner branching process model requires significantly fewer recorded outage data than empirical estimation by a factor of 10^{-1} .

This work is supported by National Natural Science Foundation of China (No. 51107040) and Fondecyt Project (No. 1110527).

Key words: cascading failure; branching process; Borel-Tanner; error analysis; probability distribution; historical outage data

中图分类号: TM71 文献标识码: A 文章编号: 1674-3415(2014)07-0023-08

0 引言

当今电力系统正在向大容量、超高压、长距离、大范围互联等方向不断发展, 由此带来良好经济效

益的同时也对电网的安全可靠运行提出了挑战。近年来国内外不断发生的大停电事故给各国造成了难以估量的经济损失和社会影响^[1-3], 引起了国际社会的广泛关注。

2001年开始, 国内外学者(涉及电力系统, 数学, 统计物理, 非线性动态等多个领域)从不同角度开始研究连锁故障的演化及干预。这些研究主要

基金项目: 国家自然科学基金资助项目(51107040);
Fondecyt项目(1110527)

着眼于: 1) 从宏观的角度研究连锁故障的特点及演化^[4-7]; 2) 连锁故障干预, 研究某一类扰动在连锁故障演化过程中的作用或识别故障链^[8-12]; 3) 对传统风险管理方法进行补充, 采用概率模型, 基于仿真数据和实际故障记录数据, 估计连锁故障停电规模的概率分布^[10-14]。例如, 文献[4-6]从自组织临界角度研究连锁故障的动力学成因, 针对连锁故障的幂律特性进行了验证; 文献[7]基于我国电网重大停电事故的统计资料, 研究了我国各大电网发生连锁故障的特征。文献[8-10]从电力系统脆弱性评估角度研究影响电网安全的关键区域和薄弱环节, 为连锁故障的预防和控制提供依据。文献[11]基于继电保护隐性故障模型和电力系统风险理论, 提出了连锁故障产生和评估的方法。文献[12]通过分析影响连锁故障过程的因素, 确定不同因素组合下的连锁故障关联度函数并用于连锁故障发展方向的实时预测。文献[13-14]采用集群模型, 分别基于实际电网历史故障记录数据和仿真数据, 采用概率方法估计连锁故障的概率分布以用于风险管理。文献[15]提出了基于分支过程理论的波雷坦尔模型对电网进行连锁故障分析, 通过和实际故障数据拟合比较, 验证了这种方法的有效性。文献[16]将文献[15]中不变的分支系数用变化的分支系数代替, 给出了改进分支过程模型的连锁故障风险分析方法。上述研究均获得了有益的研究成果。

对连锁故障进行分析和风险评估需要确定停电规模的概率分布。由上述研究模型估计的大停电事故规模的概率分布都存在不能由独立故障产生的幂率区域, 这个结果与一些国家和地区的历史故障统计数据观察所得结论相一致。针对历史故障统计数据计算, 一直是传统电力系统可靠性评估的重要组成部分。然而统计计算需要大量的样本数据, 尤其是计算停电规模的概率分布, 则必须使用多年的历史故障数据, 限制了其在年度风险管理方面的应用。

不同电网在拓扑结构、电网规模、装机容量、内在关联度、负荷水平以及其他一些物理/技术特征等方面存在很大区别。同样, 不同电网间电力系统或电力市场的规划和运行也有所不同, 尤其是国家与国家之间, 运行机制不同。不同来源的故障记录数据, 反映了不同系统的充裕性和安全性, 其分析显示各电网的可靠性特征, 既有不同, 又都显示出相似的故障规模概率分布特征^[7,13,15-16]。不同的电网, 适用的概率方法有可能不尽相同。

本文拟在文献[15]基础上, 通过对某区域南网 14 年的历史故障统计数据进行分析, 比较两种分支

过程模型—广义泊松模型和波雷坦尔模型, 以及另外两种被普遍使用的模型—泊松模型和幂率模型。最后, 通过改变历史故障记录数据覆盖时段分析利用短期故障记录数据及波雷坦尔模型估计连锁故障概率分布的精度, 研究其应用于实际电网的可能性。

1 历史故障数据特点分析

大停电事故的普遍特征之一是输电线路的相继故障, 因此, 在一定程度上, 可以利用线路故障数目衡量大停电事故的严重程度并表征大停电事故的发展过程。因此, 本文只针对所研究的目标电网的线路故障进行分析, 而不考虑其他类型故障, 例如发电机故障, 变压器故障等。

故障数据给出了目标电网 1997 年至 2011 年初约 14 年共 762 个线路故障记录数据。其中, 110 kV 及以下故障所占比例为 3.8%, 96.2% 的故障线路电压等级在 220 kV 及以上。由于 110 kV 及以下线路故障数所占比例较小, 且某些 110 kV 线路故障未提供准确的时标信息, 因此, 本文分析中仅针对 220 kV 及以上的线路故障数据, 共 733 起故障。数据中线路故障存在不同故障形式, 包括三相短路、单相短路, 自动重合闸重合于故障线路后的加速跳闸故障。由于高电压等级线路发生故障对电网的危害较大, 同样, 三相接地故障对系统的危害高于单相接地故障, 则高危害的线路故障更可能引发其他线路上发生的后续故障。由于本文分析针对初始故障引发的后续多起故障这样一个故障的传播过程, 因此, 在故障处理中可以忽略故障电压等级及故障形式的区别, 忽略线路故障的具体原因(线路断开、母线故障、继电保护误动和拒动等), 而不同危害程度的故障对电网的影响在故障的传播过程中体现。

针对连锁故障分阶段传播的特点, 本研究根据线路故障发生的时间将不同的故障归至不同连锁故障阶段。本研究假设, 调度员的调度行为一般发生在 1 h 以内, 据此可以认为故障间隔时间在 1 h 内的相邻故障属于同一连锁故障; 而故障的暂态或自动重合闸的动作一般在 1 min 内结束, 据此可以认为故障间隔时间在 1 min 内的相邻故障属于同一连锁故障的同一阶段^[15]。第一阶段发生的故障成为初始故障。根据这种方法, 733 个故障被细分成 459 个连锁故障。表 1 给出了这些连锁故障各个阶段的故障总数。从表 1 中可以看出故障传播阶段数最多达到 14。图 1 图示了按上述划分方法进行阶段和连锁过程分析后的结果。纵坐标表示所有 459 个连锁故障, 横坐标表示从该连锁故障开始(初始故障)至连锁故障结束的时间。图中每一点代表一个线路

故障, 纵坐标相同的点代表该连锁故障中的每一个子故障。对于同一时间发生的多个故障, 为了在图中显示, 将其横坐标进行了微小位移。从图 1 中可以很清楚地看出故障分阶段的聚集效果。

表 1 连锁故障各阶段故障数

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

阶段	故障数	阶段	故障数
Z_0	556	Z_8	3
Z_1	83	Z_9	3
Z_2	31	Z_{10}	3
Z_3	20	Z_{11}	2
Z_4	14	Z_{12}	2
Z_5	6	Z_{13}	1
Z_6	5	Z_{14}	1
Z_7	3	Z_{15}	0

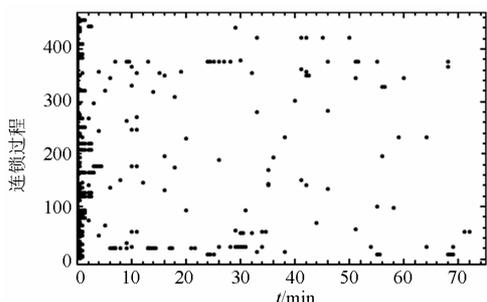


图 1 故障连锁阶段映射结果

Fig. 1 Clustering of outages in stages

表 2 根据实际数据所得线路开断数统计

Table 2 Statistics on the number of transmission line lost in the initial failure and a cascade

跳闸线路数目	初始故障次数	初始故障概率	连锁故障次数	连锁故障概率
1	402	0.875 8	341	0.742 9
2	35	0.076 25	62	0.135 1
3	13	0.028 32	27	0.058 82
4	5	0.010 89	10	0.021 79
5	0	0	4	0.008 715
6	3	0.006 536	5	0.010 89
7	1	0.002 179	5	0.010 89
8	0	0	1	0.002 179
9	0	0	1	0.002 179
10	0	0	1	0.002 179
16	0	0	1	0.002 179
19	0	0	1	0.002 179

以一个连锁故障中, 以及初始故障中发生的跳闸线路的数目衡量故障规模, 对 459 个连锁故障进行了分析。表 2 给出了分析结果。表 2 中列出了不同规模的初始故障和连锁故障的次数, 以及不同规

模的初始故障或连锁故障的发生概率。从表 2 中可知, 初始故障中跳闸线路数最大达到了 7 条, 一个连锁故障中, 跳闸线路数最大达到了 19 条; 而故障规模在跳闸线路 7 条及以下的连锁故障的概率占到约 99%, 也说明采用传统分析方法很难准确估计大规模故障的概率。由表 2 所得初始故障和总故障概率密度分布如图 2 所示。可以看出双对数坐标下, 该电网连锁故障规模的概率分布具有幂率尾现象, 这也与文献[15]中的结果相同(文献[15]采用的是同一电网 1997~2006 年初 9 年的故障数据)。

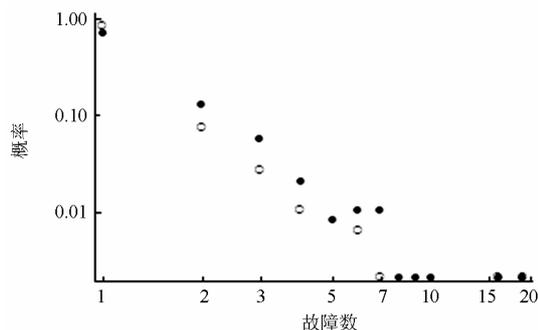


图 2 双对数坐标下初始故障(空心圆形)和总故障(实心圆形)概率分布

Fig. 2 Log-log plots of probability distribution of initial (hollow circle) and total (solid circle) line outages

2 分支过程模型

本节介绍用于估计连锁故障规模概率分布的分支过程模型。分支过程模型并不直接揭示故障传播的物理过程和机理, 而是通过实际故障数据估计分支过程模型的参数, 进而根据这些参数估计电力系统发生的故障规模的概率分布。

分支过程为随机过程, 描述的是一组粒子的分裂或灭亡的过程。根据分支过程的特点, 可以将连锁故障演化过程作如下描述: 初始有 $Z_0 (Z_0 > 0)$ 条线路断开, 由于这些线路断开, 使电网可靠性下降, 从而导致下一阶段 Z_1 条线路断开; 继续又在下一个阶段产生 Z_2 条线路断开…。如此传播, 直到某一阶段没有发生故障, 则该阶段表示这个连锁故障过程的结束。

2.1 分支概率

分支过程模型的重要参数之一是分支概率 λ , 即由任何一阶段的一个故障在下一阶段产生 r 个故障 ($r=0,1,2,\dots$) 的概率。由于所有阶段的故障都有可能在下阶段产生新的故障, 而所有从第二阶段到最后阶段的故障分别由它们上一阶段的故障所产生。因此, 我们可采用以下方法估计 λ 。

假设一共存在 J 个连锁故障, 令 $Z_k^{(j)}$ ($j=1,2,\dots,J; k=0,1,\dots$) 代表第 j 个连锁故障的第 k 个阶段的故障数目, 则分支概率的估计值 $\hat{\lambda}$ 为^[17]

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

其中: J 为总故障数; $N(i)$ 为线路故障数目不为 0 的最大阶段数。

将表 1 数据代入式 (1) 计算, 可得

$$\hat{\lambda} = \frac{Z_1 + Z_2 + \dots + Z_{14}}{Z_0 + Z_1 + \dots + Z_{13}} \approx 0.24 \quad (2)$$

当 $\hat{\lambda}$ 取值大于 1 时, 表示连锁故障将会不断发展, 直至连锁故障规模达到系统规定的最大规模; 当 $\hat{\lambda}$ 取值小于 1 时, 故障规模逐渐衰减到 0, 连锁故障停止。根据计算所得 $\hat{\lambda} < 1$ 表示所有连锁故障最终都会趋于结束。

2.2 连锁故障规模概率分布模型

分支过程模型中的另外一个重要的组成部分是初始故障的概率分布。如果初始故障分布假设为泊松分布, 则连锁故障概率分布为广义泊松分布; 如果初始故障为任意分布, 则连锁故障概率分布为波雷-坦尔分布^[18]。

(1) 广义泊松模型

若初始故障分布服从参数为 θ 的泊松分布且不考虑初始故障为 0 的情况, 则初始故障概率分布为

$$P[Z_0 = r] = \frac{e^{-\theta} \theta^r}{(1 - e^{-\theta}) r!}, r = 1, 2, \dots \quad (3)$$

其中, r 为故障线路数目。其均值 \bar{Z}_0 为

$$\bar{Z}_0 = \hat{\theta} / (1 - e^{-\hat{\theta}}) \quad (4)$$

这种情况下连锁故障规模服从广义泊松分布, 如式(5)所示。

$$P[Z = r] = \hat{\theta} (r \hat{\lambda} + \hat{\theta})^{r-1} \frac{e^{-r \hat{\lambda} - \hat{\theta}}}{(1 - e^{-\hat{\theta}}) r!}, r = 1, 2, \dots \quad (5)$$

其中, 由式 (2) 可得 $\hat{\lambda}$ 取值。根据表 2 可知 $\bar{Z}_0 \approx 1.211$, 于是由式 (4) 可得 $\hat{\theta} \approx 0.3965$ 。

(2) 波雷-坦尔模型

若初始故障具有任意分布, 则连锁故障规模服从波雷-坦尔分布。

$$P[Z = r] = \sum_{z_0=1}^r P[Z_0 = z_0] z_0 \hat{\lambda} (r \hat{\lambda})^{r-z_0-1} \frac{e^{-r \hat{\lambda}}}{(r - z_0)!} \quad (6)$$

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

其中, $P[Z_0 = z_0]$ 可由实际故障数据计算所得, 参数 $\hat{\lambda}$ 取值由式 (2) 计算得到。

3 结果分析

本节以电网历史故障数据为样本数据, 分别采用基于分支过程的概率模型(包括广义泊松分布模型和波雷-坦尔模型), 传统泊松模型和幂率模型(参见附录 A) 对连锁故障概率分布进行估计, 并进行比较, 分析各模型的适用性。

3.1 故障规模概率估计

将参数 $\hat{\theta}$ 、 $\hat{\lambda}$ 取值和不同故障线路数 r 值分别代入式 (5)、式 (6)、式 (A5)、式 (A7) 进行计算, 可得到不同模型估计所得概率分布。图 3 显示了不同概率模型估计所得结果。

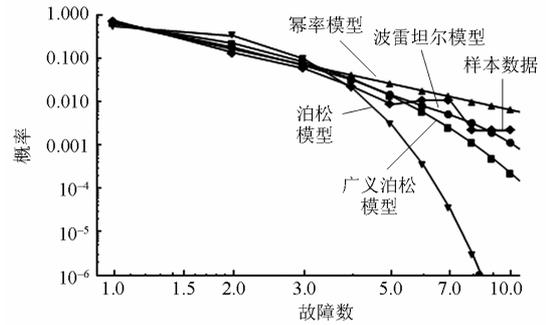


图 3 双对数坐标下不同模型概率分布

Fig. 3 Log-log plots of probability distribution in different models

从图 3 可知, 小规模故障(故障线路数少于 5 条)时, 四种模型都能很好地估计故障概率。而随着故障规模的增大, 波雷坦尔模型与实际故障数据的拟合程度则明显优于其他三种模型。其中, 幂率模型高估了发生大停电事件的概率, 实际问题中, 不管是电力系统的停电故障数据^[19]还是其他领域内的数据(如国家 GDP、城市人口等)^[20], 其幂率分布特征一般呈现在概率密度函数曲线的尾部, 以停电故障为例, 传统可靠性研究认为故障概率分布为指数分布, 而实际故障数据及研究^[19]表明, 发生故障的概率分布呈现部分幂率, 即大型故障呈现幂率(分布的尾部)。同时, 有研究表明系统处于自组织临界状态时, 其故障概率分布呈现幂率。然而系统临界与否或临界点不同, 其整个分布可能呈现先指数后幂率或不同斜率的幂率分布。因此, 仅用幂率模型来描述连锁故障概率分布可能会存在一定的误差; 泊松模型和基于分支过程的广义泊松模型都低估了其发生的概率。其中, 泊松模型估计所得概率分布与实际样本数据概率分布偏离严重, 故障线

路数超过5的故障概率为实际概率的 10^{-2} 到 10^{-5} 倍。这是由于泊松模型假设每条线路断开为相互独立事件, 而实际情况中, 当连锁故障发展到后期阶段, 系统运行状态趋于恶化, 有些故障发生概率大大增加, 例如保护的隐性故障、线路由于过负荷发热弧垂增大等, 并且一些小的故障很可能引发一系列的潮流波动, 从而导致更多故障发生。基于分支过程的广义泊松模型相较于幂率模型和泊松模型, 与实际数据的拟合程度更高。它与波雷坦尔模型相同, 都是以分支过程模型为核心, 现阶段研究表明, 以分支概率体现连锁故障分阶段逐渐传播的分支过程模型对连锁故障演化描述最好; 它与波雷坦尔模型的区别在于初始故障的分布, 广义泊松模型假设初始分布为泊松分布, 而此假设仍需要大量实际数据验证。波雷坦尔模型假设初始故障为任意分布, 回避了初始故障分布对该模型精度的限制。为了更加直观地显示出波雷坦尔模型较其他模型的优越性, 下一小节将通过卡方拟合优度检验对这些模型的拟合度进行比较。

3.2 拟合优度检验

卡方拟合优度检验^[13]是采用数学方法判断概

率模型得到的理论分布和由实际数据得到的样本分布的相似程度的方法之一。

卡方拟合优度检验方法简述如下:

假设 X 为取值 $1, 2, \dots, k$ 的离散分布, 且 $P(X=i)=p_i, 1 \leq i \leq k$ 。对 X 进行 n 次观测, $X_i (i=1, 2, \dots, k)$ 为 n 次观测中 X 取 i 的次数。则可以通过统计量 χ^2 来比较 X 为各取值的期望频数与观测频数的差异。

$$\chi^2 = \sum_{i=1}^k \frac{(X_i - np_i)^2}{np_i} \quad (7)$$

当 $n \rightarrow \infty$ 时, χ^2 的极限分布服从自由度 $k-1$ 的 χ^2 分布。从式 (7) 可知, χ^2 越大, 样本数据偏离概率模型越大。上述结论成立的条件是模型分布为多项式分布且所有的期望频数 np_i 的均值应大于 5, 因此我们将样本空间分为互不相交的 5 个子空间: $M_1=\{1\}$, $M_2=\{2\}$, $M_3=\{3\}$, $M_4=\{4\}$, $M_5=\{5, 6, \dots\}$ 。其中, $M_i=\{i\} (i=1, 2, 3, 4)$ 表示该空间包含故障数取值为 i 的情况; $M_5=\{5, 6, \dots\}$ 表示该空间包含的故障数为 $5, 6, \dots$ 。经过上述空间划分后, 总故障数位于各子空间的连锁故障发生次数 $X_i (i=1, 2, 3, 4, 5)$ 服从多项式分布。卡方检验结果如表 3 所示。

表 3 不同模型的卡方值

Table 3 χ^2 test results for all models

r	X_i	波雷坦尔		广义泊松		幂率		传统泊松	
		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
自由度		3		2		3		3	
χ^2		9.31		34.07		22.87		292.4	

当概率模型中有 s 个需要估计的参数时, χ^2 分布的自由度变为 $k-s-1$ 。从表 3 可知, 由波雷坦尔模型计算出的 χ^2 值明显小于其他模型, 即波雷坦尔模型和样本数据的拟合程度最好。令 $\chi_\alpha^2(m)$ 为显著性水平 α , 自由度 m 下的卡方值, 则当 α 为 0.01 时, 波雷坦尔模型的 $\chi^2 < \chi_{0.01}^2(3) = 11.34$, 即在显著性水平为 0.01 下, 波雷坦尔模型能很好描述实际故障数据的分布。

4 波雷-坦尔模型的适用性分析

由于不同概率模型对不同电网和故障数据源的适应性不同, 本节拟通过分析针对同一电网, 当

采用的历史故障记录数据所覆盖的时段发生变化, 即样本数量发生变化时, 波雷-坦尔模型的适应性。本小节将分别采用 14 年故障记录数据和 1 年故障记录数据作为样本数据讨论由波雷坦尔模型估计得到的 r 条线路发生故障 (即故障规模为 r) 的概率 \hat{F}_r 的精度。

4.1 估计误差分析方法

当给定某一样本值 F_r 时, 可通过计算概率 $P\{F_r/2 \leq \hat{F}_r \leq 2F_r\}$ 评估估计值 \hat{F}_r 的精度, 其中系数 2 为经验取值。

假设 X 是具有均值 μ , 标准差 σ 的高斯随机变量, 令 $\kappa(X)$ 为变异系数, 满足式 (8) 的关系。

$$\kappa(X) = \mu / \sigma \quad (8)$$

则当置信度取值为 90%，即

$$P(\mu / 2 \leq X \leq 2\mu) = P(-\kappa / 2 \leq \frac{X - \mu}{\sigma} \leq \kappa) \geq 0.90 \text{ 时,}$$

可以得到 $\kappa(X) \geq 2.61$;

当置信度取值为 95% 时，可以得到 $\kappa(X) \geq 3.30$ 。

因此，可以通过计算变异系数是否大于 2.61 (或 3.30)，判定概率 $P\{F_r / 2 \leq \hat{F}_r \leq 2F_r\}$ 是否大于等于 90% (或 95%)。如果变异系数大于 2.61 (或 3.30)，则说明由波雷-坦尔模型估计得到的 r 条线路发生故障 (即故障规模为 r) 的概率 \hat{F}_r 精确的置信度为 90% (或 95%)。

4.2 模型的适用性分析

由式 (6) 可知，故障规模为 r 的概率 F_r 由 λ 的取值确定，通过对模型 (6) 线性化可得 F_r 的标准差为

$$\sigma(F_r) = |D_\lambda F_r|_{\lambda=0.24} \times \sigma(\lambda) \quad (9)$$

式中：第一项表示 λ 取 0.24 时， F_r 对 λ 求导所得绝对值；第二项标准差 $\sigma(\lambda)$ 可由式 (10) 计算^[16]。

$$\sigma^2(\lambda) = \frac{(1-\lambda)(1-e^{-\lambda})}{J(1-\lambda^{14})} \quad (10)$$

其中， J 表示实际故障数据中连锁故障总数。则 F_r 的变异系数为 $\kappa(F_r) = F_r / \sigma(F_r)$ ，其中 $\kappa(F_r)$ 为基于所有 14 年故障数据所得的变异系数，相应的基于 1 年故障数据所得变异系数为 $\kappa(F_r) / \sqrt{14}$ 。

计算结果显示：以 2 为系数，置信度为 95% 时，采用波雷坦尔模型基于 14 年故障数据能精确估计故障概率 F_r 直到 $r=13$ 。其中， $\kappa(F_1)=55.5$ ，随着 r 不断增大， $\kappa(F_r)$ 呈逐渐减小趋势，并且 $\kappa(F_{13})=3.45$ ， $\kappa(F_{14})=2.95$ 。置信度为 95% 时，采用此模型基于 1 年故障数据能精确估计故障概率 F_r 直到 $r=4$ ；而当置信度为 90% 时，能精确估计故障概率的线路故障数达到 $r=8$ 。这说明根据 1 年故障数据，采用波雷坦尔模型也能较好估计电网总故障数的概率分布，并且可以通过增加样本数据覆盖时段来提高估计精度。

表 4 左边给出了基于 1 年故障数据，对应 $r=1, 2, \dots, 10$ 的变异系数。表 4 右边给出了由实际 14 年故障数据计算所得概率分布 p_r 及其标准差 $\sigma(p_r)$ 及变异系数 $\kappa(p_r)$ 。

表 4 故障概率分布及其变异系数

Table 4 Distribution of line outages and their coefficients of variation

r	F_r	$\sigma(F_r)$	$\frac{\kappa(F_r)}{\sqrt{14}}$	p_r	$\sigma(p_r)$	$\kappa(p_r)$
1	0.688	0.012 4	14.8	0.743	0.020 4	36.4
2	0.178	0.003 37	14.1	0.135	0.016 0	8.44
3	0.068 8	0.003 17	5.80	0.058 8	0.011 0	5.35
4	0.031 3	0.002 14	3.90	0.021 8	0.006 81	3.20
5	0.014 2	0.001 41	2.70	0.008 71	0.004 34	2.01
6	0.007 94	0.000 787	2.70	0.010 9	0.004 85	2.25
7	0.005 05	0.000 429	3.15	0.010 9	0.004 85	2.25
8	0.003 18	0.000 296	2.86	0.002 18	0.002 18	1.00
9	0.001 91	0.000 227	2.25	0.002 18	0.002 18	1.00
10	0.001 10	0.000 170	1.73	0.002 18	0.002 18	1.00

从表 4 中可知，当 $r \geq 2$ 时，基于 1 年故障数据的 $\kappa(F_r) / \sqrt{14}$ 超过了基于 14 年故障数据的 $\kappa(p_r)$ 。这说明在同等精度要求下，采用波雷-坦尔模型估计故障分布所需样本数据比直接根据实际故障数据计算所得概率分布所需样本数据整整降低一个数量级。这一特点极大地提高了连锁故障概率估计的效率，体现了波雷-坦尔模型应用于实际电网的可能性。管理部门可以针对每年的故障记录数据，采用波雷-坦尔模型估计电网发生连锁故障的概率分布，与往年的故障规模概率分布进行比较，为制定下一年度

降低电网发生大型停电故障的风险的措施提供参考。尽管采用波雷-坦尔模型估计连锁故障发生概率的方法相较于其他传统模型估计方法具有明显的优点，但对于大规模故障，存在大量实际故障数据时这种方法估计概率精度才能保持一定置信度。为克服这个缺点，可以通过计算每个阶段的分支系数来代替恒定不变的分支系数，并且由于系统线路断开数目越多，系统越脆弱，在后期阶段的分支系数计算中可引入一定的权重系数。往后将在这方面进行更深入的研究来完善本文提到的模型。

5 结语

本文以某区域电网 14 年的线路故障数据为样本数据, 比较研究了传统泊松、幂率、广义泊松和波雷-坦尔四种概率模型用于电网连锁故障概率分布估计的适用性。本文通过分析保护及自动装置以及调度行为的特点, 对电网 14 年的故障记录数据进行了分析及预处理, 并将其用于分支过程模型参数的计算。分析结果表明, 与传统概率模型相比较, 波雷-坦尔模型能够更好地估计故障规模的概率分布, 与实际样本数据分布具有很高的拟合度。而针对同一组样本数据, 幂率模型高估了发生较大规模故障的概率, 泊松模型和广义泊松分支过程模型低估了其发生的概率。误差分析结果表明, 在相同置信度要求下, 基于波雷-坦尔模型估计故障概率分布所需样本数据比直接根据实际故障数据计算所得概率分布所需样本数据降低一个数量级, 进一步验证了波雷-坦尔分支过程模型应用于实际电网风险管理的有效性和可能性。

感谢 Ian Dobson 教授提供的分支过程模型及仿真程序, 同时感谢相关电力公司提供用于本文研究的故障数据记录!

附录 A

A.1 传统泊松模型

设随机变量 $T \in \{0, 1\}$, 其中 $T=1$ 代表一条独立线路断开且其发生概率为 $P(T=1)=p$, 则任何一条线路断开概率可表示为

$$P(T=t|p) = p^t(1-p)^{1-t}, t=0, 1; 0 \leq p \leq 1 \quad (\text{A1})$$

假设目标电网总线路数为 N , 每条线路断开概率为 p 且每条线路断开都是相互独立事件, 则总线路断开数 Z 服从二项式分布。

$$P[Z=r] = C_N^r p^r (1-p)^{N-r}, r=0, 1, 2, \dots, N \quad (\text{A2})$$

通常, 式 (A2) 中 N 很大且 p 较小, 于是式 (A2) 可近似表示为参数为 $\lambda_{\text{con}}=Np$ 的泊松分布。

$$P[Z=r] = C_N^r p^r (1-p)^{N-r} \approx \frac{e^{-\lambda_{\text{con}}} \lambda_{\text{con}}^r}{r!}, r=0, 1, 2, \dots, N \quad (\text{A3})$$

本文采用最大似然估计法计算参数 λ_{con} 。若总体 X 为离散型, 假设其分布律为 $P(X=x)=p(x; \eta)$, η 为待估参数。则事件 $\{X_1=x_1, X_2=x_2, \dots, X_n=x_n\}$ 的联合分布律为

$$L(\eta) = L(x_1, \dots, x_n; \eta) = \prod_{i=1}^n p(x_i; \eta) \quad (\text{A4})$$

使式 (A4) 取得最大值的参数值 $\hat{\eta}$ 称为最大似然估计量, 于是我们可通过求导等方法求出参数 $\hat{\eta}$ 。

根据表 2 及上述最大似然估计法, 可得 $\lambda_{\text{con}} \approx 0.5970$, 则泊松模型计算故障概率表达式为

$$P[Z=r] = e^{-0.5970} 0.5970^{r-1} / (r-1)!, r=1, 2, \dots \quad (\text{A5})$$

A.2 幂率模型

当总线路断开数 Z 服从幂率分布时, 其概率分布函数为

$$P[Z=r|q] = r^{-q} / \sum r^{-q}, q > 0; r=1, 2, \dots \quad (\text{A6})$$

在双对数坐标下作出 $P[Z=r]$ 与 r 的关系图时, 可得一条斜率为 $-q/\sum r^{-q}$ 的直线, 说明其概率分布服从幂率分布。同样采用最大似然估计法, 根据表 2 可得 $\hat{q} \approx 2.0$, 于是可得幂率模型计算故障概率表达式为

$$P[Z=r] = r^{-2.0} / \sum_{k=1}^{\infty} k^{-2.0} = 0.645r^{-2.0}, r=1, 2, \dots \quad (\text{A7})$$

参考文献

- [1] 印永华, 郭剑波, 赵建军, 等. 美加“8.14”大停电事故初步分析及应吸取的教训[J]. 电网技术, 2003, 27(10): 8-11, 16.
YIN Yong-hua, GUO Jian-bo, ZHAO Jian-jun, et al. Preliminary analysis of large scale blackout in interconnected North America power system grid on August 14 and lessons to be drawn[J]. Power System Technology, 2003, 27(10): 8-11, 16.
- [2] 葛睿, 董昱, 吕跃春. 欧洲“11.4”大停电事故分析及对我国电网运行工作的启示[J]. 电网技术, 2007, 31(3): 1-6.
GE Rui, DONG Yu, Lü Yue-chun. Analysis of large-scale blackout in UCTE power grid and lessons to be drawn to power grid operation in China[J]. Power System Technology, 2007, 31(3): 1-6.
- [3] 董新洲, 曹润彬, 王宾, 等. 印度大停电与继电保护的三大功能[J]. 电力系统保护与控制, 2013, 41(2): 19-25.
DONG Xin-zhou, CAO Run-bin, WANG Bin, et al. India blackout and three functions of protective relay[J]. Power System Protection and Control, 2013, 41(2): 19-25.
- [4] DOBSON I, CARRERAS B A, LYNCH V E, et al. An initial model for complex dynamics in electric power system blackouts[C] // Proceedings of the 34th Hawaii International Conference on System Sciences, January 3-6, 2001, Maui, Hawaii, USA: 710-718.
- [5] REN H, DOBSON I, CARRERAS B A. Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid[J]. IEEE Trans on Power Systems, 2008, 23(3): 1217-1225.
- [6] 梅生伟, 何飞, 张雪敏, 等. 一种改进的 OPA 模型及大停电风险评估[J]. 电力系统自动化, 2008, 32(13):

- 1-5.
MEI Sheng-wei, HE Fei, ZHANG Xue-min, et al. An improved OPA model and the evaluation of blackout risk[J]. Automation of Electric Power Systems, 2008, 32(13): 1-5.
- [7] 于群, 郭剑波. 中国电网停电事故统计与自组织临界性特征[J]. 电力系统自动化, 2006, 30(2): 16-21.
YU Qun, GUO Jian-bo. Statistics and self-organized criticality characters of blackouts in China Electric Power Systems[J]. Automation of Electric Power Systems, 2006, 30(2): 16-21.
- [8] 舒征宇, 马俊民, 邓长虹, 等. 小世界电网的故障传播特性分析与运行极限匹配模型研究[J]. 电力系统保护与控制, 2013, 41(7): 29-35.
SHU Zheng-yu, MA Jun-min, DENG Chang-hong, et al. Analysis of fault propagation characteristics of small-world power system and research of robust match model[J]. Power System Protection and Control, 2013, 41(7): 29-35.
- [9] 魏震波, 刘俊勇, 朱国俊, 等. 基于可靠性加权拓扑模型下的电网脆弱性评估模型[J]. 电工技术学报, 2010, 25(8): 131-137.
WEI Zhen-bo, LIU Jun-yong, ZHU Guo-jun, et al. Vulnerability evaluation model to power grid based on reliability-parameter-weighted topological model[J]. Transactions of China Electrotechnical Society, 2010, 25(8): 131-137.
- [10] 尹项根, 陈庆前, 王博, 等. 基于四层集合模型的复杂电力系统脆弱性评估体系[J]. 电工技术学报, 2013, 28(1): 225-233.
YIN Xiang-gen, CHEN Qing-qian, WANG Bo, et al. Vulnerability assessment system of complicated power system based on four-level set model[J]. Transactions of China Electrotechnical Society, 2013, 28(1): 225-233.
- [11] 杨明玉, 田浩, 姚万业. 基于继电保护隐性故障的电力系统连锁故障[J]. 电力系统保护与控制, 2010, 38(9): 1-5.
YANG Ming-yu, TIAN Hao, YAO Wan-ye. Analysis of power system cascading failure based on hidden failures of protective relaying[J]. Power System Protection and Control, 2010, 38(9): 1-5.
- [12] 王英英, 罗毅, 涂光瑜, 等. 电力系统连锁故障的关联模型[J]. 电工技术学报, 2012, 27(2): 204-209.
WANG Ying-ying, LUO Yi, TU Guang-yu, et al. Correlation model of cascading failures in power system [J]. Transactions of China Electrotechnical Society, 2012, 27(2): 204-209.
- [13] CHEN Q M, JIANG C W, QIU W Z, et al. Probability models for estimating the probabilities of cascading outages in high-voltage transmission network[J]. IEEE Trans on Power Systems, 2006, 21(3): 1423-1431.
- [14] 于会泉, 刘文颖, 温志伟, 等. 基于线路集群的连锁故障概率分析模型[J]. 电力系统自动化, 2010, 34(10): 29-33.
YU Hui-quan, LIU Wen-ying, WEN Zhi-wei, et al. A probabilistic assessment model for power system cascading failure based on line clusters method[J]. Automation of Electric Power Systems, 2010, 34(10): 29-33.
- [15] REN H, DOBSON I. Using transmission line outage data to estimate cascading failure propagation in an electric power system[J]. IEEE Trans on Circuits and Systems, 2008, 55(9): 927-931.
- [16] DOBSON I. Estimating the propagation and extent of cascading line outages from utility data with a branching process[J]. IEEE Trans on Power Systems, 2012, 27(4): 2146-2155.
- [17] DOBSON I, CARRERAS B A, NEWMAN D E. A loading dependent model probabilistic cascading failure[J]. Probability in the Engineering and Informational Sciences, 2005, 19(1): 15-32.
- [18] 任惠. 电力系统连锁故障风险分析[D]. 保定: 华北电力大学, 2009: 54-55.
REN Hui. Risk analysis of power system cascading blackouts[D]. Baoding: North China Electric Power University, 2009: 54-55.
- [19] CARRERAS B A, NEWMAN D E, DOBSON I, et al. Evidence for self-organized criticality in a time series of electric power system blackouts[J]. IEEE Trans on Circuits and Systems, 2004, 51(9): 1733-1740.
- [20] 胡海波, 王林. 幂率分布研究简史[J]. 物理, 2005, 34(12): 889-896.
HU Hai-bo, WANG Lin. A brief history of power law distributions[J]. Physics, 2005, 34(12): 889-896.

收稿日期: 2013-07-03; 修回日期: 2013-12-10

作者简介:

任惠(1973-), 女, 博士, 教授, 主要研究方向为电力系统风险评估, 新能源电力系统与智能电网等;

熊吉(1988-), 男, 硕士研究生, 研究方向为电力系统连锁故障风险分析; E-mail: xiongji.tian@163.com

David Watts (1973-), 男, 博士, 副教授, 主要研究方向为电力市场设计、运行及输电配电定价, 能源规划和建模, 新能源和能源效率及电力系统连锁故障风险。