

基于贝叶斯网络的短期负荷预测方法

作为电力系统经济调度中的重要内容之一,短期负荷预测的准确性对改善电网运行性能起着重要作用。由于传统预测方法具有一定的缺陷,提出了一种基于贝叶斯网络的短期负荷预测方法,在综合考虑影响负荷要素的基础上构建贝叶斯网络模型,并在已知负荷数据的基础上提出基于贝叶斯法则的学习算法,对某已知电网的大量负荷数据进行学习和预测。预测结果表明,贝叶斯网络模型较传统多元线性回归方法相比具有更高的预测准确度。

周玉^{1,2} 崔高颖^{1,2} 易永仙^{1,2} 陈霄^{1,2} 石坤³ 许高杰^{3/1} 江苏省电力公司
电力科学研究院 2.国家电网公司电能计量重点实验室 3.中国电力
科学研究院



周玉/工程师

0 引言

作为能量管理系统(EMS)的一个重要模块,短期电力负荷预测的准确性对改善电网运行性能起着重要作用^[1]。但预测准确度易受多种因素不同程度的影响。这些因素包括气象、经济发展速度和重大事件等^[2],而气象随着自然条件的变化而发生改变,同时在不同地区存在较大差异;不同地区的经济发展速度也存在差距;重大事件的发生更加是不可预料的;另外,相同因素的不同情况对预测准确度也将产生不同影响,故要实现准确地预测负荷具有较大难度。

目前的负荷预测方法主要可以分为时间序列预测法^[3]、回归分析预测法^[4]和人工智能预测法^[5]。时间序列预测法将历史电力负荷数据当作是按一定时间间隔变化的一个时间序列,建立一个能够描述此时间序列基本特征

和变化规律的数学模型,由该模型再得出能够预测未来负荷的数学表达式。时间序列法主要有自回归(AR)模型、滑动平均(MA)模型^[2]、自回归-滑动平均(ARMA)模型、累计式自回归-滑动平均(ARIMA)模型和传递函数(TF)模型等。回归分析预测法是参考历史电力负荷数据,构造具有分析功能的数学模型,实现预测未来负荷的目的。其关键特征为把影响负荷的各个因素当作自变量,把负荷当作因变量,表现出较明显的内插优势。从数学角度来看,回归分析法实际上是通过统计并分析解释变量的观测数据来确定电力负荷与解释变量之间的函数关系,进而达到负荷预测的目的。然而,负荷预测的影响因素具有种类多、突发性和不确定性等特点,导致上述负荷预测方法的准确度较低,尤其是节假日等特殊负荷^[6],且鲁棒性较差,预测效果的参考价值较低。近年来,人工智能方法得到快速发展,具有代表性的方法如人工神经网络、专家系统和支持向量机预测等方法。

电力系统在长期运行中监控系统会记录大量的数据^[7],如何从这些数据中提取规律性的知识是很有价值的研究方向。在分析历史数据的基础上,采用贝叶斯网络^[8-9]构建负荷预测模型,通过确定贝叶斯网络的结构明确影响负荷的主要因素,同时根据网络关系推导未来时刻的负荷预测值。目前,贝叶斯网络预测已应用多个工业领域^[10-13],但利用贝叶斯网络进行负荷预测的研究还很少见。

关键词/Keywords

负荷预测 ·
贝叶斯网络 ·
多元线性回归 ·
预测准确度 ·

1 数学模型

1.1 贝叶斯网络

贝叶斯网络是贝叶斯方法与图形理论的有机结合，是为了解决不确定性和不完整性问题而提出的一种计算方法，也是一种以贝叶斯公式为基础的概率网络。贝叶斯网络具有以下特性：本身是一种不定性因果关联模型；具有强大的不确定性问题处理能力；能够有效地进行多渠道信息表达及融合。故考虑将贝叶斯网络应用于电力负荷预测领域，从而为其提供一种新的解决方案。

贝叶斯网络是基于概率推理的数学模型，而概率推理即根据相关变量的信息来得到所需概率信息的过程。设 $A_1, A_2, \dots, A_n$ 为一组完备事件，且满足下面 3 个条件，即

$$\begin{cases} \bigcup_{i=1}^n A_i = \Omega \\ A_i A_j = \varnothing, \quad i \neq j \\ P(A_i) > 0 \end{cases} \quad (1)$$

则贝叶斯公式可表示为

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)} \quad (2)$$

设事件集合 $\{A_1, A_2, \dots, A_n\}$ 对应 n 个随机变量 $X=\{X_1, X_2, \dots, X_n\}$ ，其联合概率分布即为贝叶斯网络 $A=\left[A_s, A_p\right]$ 。其中 A_s 为有向无环图 (DAG)， A_p 为 n 组条件概率分布 (CPD)。通常，DAG 中的节点表示随机变量，连接两个节点的箭头代表这两个随机变量是具有因果关系或是非条件独立的，它们之间会产生一个条件概率值，而节点中变量间若没有箭头相互连接的情况即代表这两个随机变量是条件独立的。图 1 为贝叶斯网络结构的示意图，图 1 中共有 6 个节点，每个箭头则代表这两个节点之间存在一个条件概率。

根据概率论可知，给定 A_s 条件下， X 的联合概率分布可表示为

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \cap_{X_i}) \quad (3)$$

式中， $\cap_{X_i}$ 为节点 X_i 的父节点集，如果 X_i 无父节点，则 $\cap_{X_i} = \varnothing$ 。

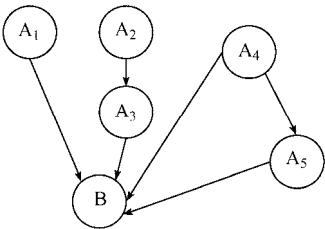


图 1 贝叶斯网络结构示意图

1.2 贝叶斯网络的学习

贝叶斯网络的学习包括贝叶斯网络结构 (DAG) 的学习和结构参数 (CPD) 的学习^[14-15]。

DAG 学习就是为了搜索与样本数据 I 具有最佳匹配度的贝叶斯网络结构。使用基于评分的算法，即将贝叶斯网络当作已包含变量之间联合概率分布的结构。

评价匹配度高低就需要给定评分函数——贝叶斯后验概率，再使用有启发法则（如断定某些节点之间必然有箭头连接，减少搜索时间）的带随机重启（为解决贪心算法可能落入局部最优的问题）的贪心算法搜索网络结构，算法流程图如图 2 所示。已有学者证实，在相同搜索时间条件下，该算法可以搜索到比模拟退火和最好优先方法的匹配度更高的结构。

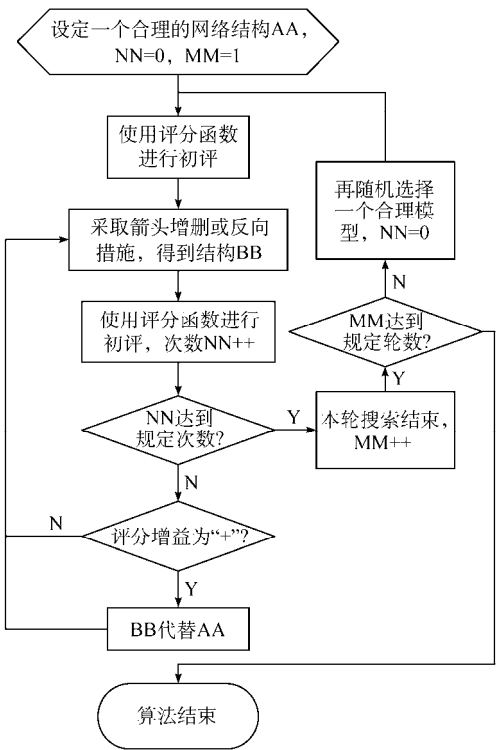


图 2 带随机重启的贪心算法流程图

在 I 已知时, 贝叶斯网络结构为 G 的条件概率 (即评分函数) 可表示为

$$P(G^h|I) = P(G^h)P(I|G^h)/P(I) \quad (4)$$

式中, G^h 为 X 的联合概率分布假设能够按照 G 分解; $P(I)$ 为不受 G^h 影响的标准化常数, 可忽略; $P(G^h)$ 为搜索到的结构是 G 的概率, 其值由用户设定。故式 (4) 中仅有 $P(I|G^h)$ 未知。

采用文献[12]中求解 $P(I|G^h)$ 的解析式, 即

$$P(I|G^h) = \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{\Gamma(a_{ij})}{\Gamma(a_{ij} + N_{ij})} \cdot \prod_{k=1}^{n_i} \frac{\Gamma(a_{ijk} + N_{ijk})}{\Gamma(a_{ijk})} \quad (5)$$

式中, n 为 I 中每项记录所包含的变量数; q_i 为第 i 个变量的父节点个数; N_{ijk} 为 I 中第 i 个变量的父节点集是第 j 个且其取值为第 k 个值的实例个数; N_{ij} 为 N_{ijk} 对 k 求和; a_{ij} 为超参数。

CPD 学习需根据实际数据进行计算, 具体过程可参考文献[12]和文献[14], 此处不再赘述。

2 负荷预测

根据贝叶斯理论和历史数据, 筛选出了贝叶斯网络中的 5 个关联参数, 分别为降雨量、季节、温度、前一天负荷和节假日。

影响负荷量的变量集 $D=\{R,S,T,F,H\}$, 各个变量的定义为: D 为负荷预测量; R 为降雨量; S 为季节; T 为温度; F 为前一天负荷; H 为节假日。

根据某地区一年负荷量数据进行相关性分析建立如图 3 所示的贝叶斯网络模型的网络结构, 且 D 中的每个变量分别与该图中的某个节点相对应。

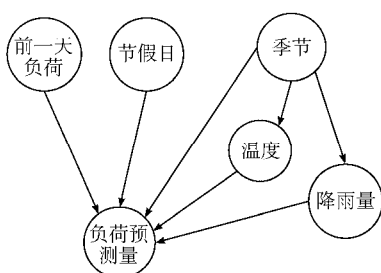


图 3 负荷预测模型网络结构

从图 3 网络结构能够看出, 不论是前一天负荷或节假日, 还是季节、温度或降雨量, 每个变量的状态均会对负荷预测量产生一定影响。求解 D 在上述 5 个关联参数联合分布下的条件概率, 即计算概率 $P(D|R,S,T,F,H)$, 也就是一个贝叶斯推理的过程。

根据式 (2) 可得

$$P(D|R,S,T,F,H) = \frac{P(R,S,T,F,H|D)P(D)}{P(R,S,T,F,H)} \quad (6)$$

$$= P(R,S,T,F,H|D)P(D)$$

根据条件独立原则可知

$$P(R,S,T,F,H|D) = P(F|D)P(H|D)P(R,S,T|D) \quad (7)$$

进一步转化为

$$P(D|R,S,T,F,H) = P(H|D)P(F|D)P(R,S,T|D)P(D) \quad (8)$$

由式 (8) 可知, 要求解 $P(D|R,S,T,F,H)$ 的值, 需得到某些先验概率和条件概率, 通过对历史数据的统计分析可得

$$P(R,S,T,F,H|D) = P(F)P(H)P(T|S)P(R|S)P(H|D)P(F|D)P(R,S,T|D)P(D) \quad (9)$$

将式 (8) 代入式 (9) 可得图 2 中各个变量的联合概率分布, 即

$$P(R,S,T,F,H|D) = P(F)P(H)P(T|S)P(R|S)P(H|D)P(F|D)P(R,S,T|D)P(D) \quad (10)$$

根据式 (10) 可通过反复使用贝叶斯公式和乘积与求和公式求解出网络结构中任意想知道的概率。

以上推导解决了概率计算问题, 但负荷预测需要得出具体的数值。因此需要量化各个变量的状态, 设某一关联参数的物理范围为 L_{lower} 和 L_{upper} , 将该物理范围等分为 m 个区间, 则第 t 个区间可表示为

$$L_t = \frac{L_{upper} - L_{lower}}{m}t \quad (11)$$

式中， $t \in [1, m]$ 且为整数，每个不同的关联参数可选择不同的区间数进行量化， m 的值越大，模型越精确计算量就越大，建立网络时需要的历史数据也就多。例如可将符合预测量的范围等分为 100 份，每个区间相当于标么值 0.01 pu，这样预测准确度即为 0.01 pu，由于为应对负荷变化而预留的备用容量通常为最高负荷的 1%~5%，而 0.01 pu 的准确度相当于误差为 1%， $m=100$ 的设定符合电力系统的实际。当然 m 的取值还可以更大，从而得到解更精确，但需要注意历史数据是有限的， m 越大相当于关联参数的状态也就越多，就需要有海量的数据才能满足统计分析的需要。随着数据存储技术和电力系统监控手段的提高以及大数据分析能力的提升，本文模型的准确度还可以进一步提升。将参数状态化后，可根据 5 个关联参数的状态预测负荷处于哪个状态，即负荷处于哪个区间从而实现负荷预测。

3 算例分析

以某地区 2010 年 1 月至 2011 年 12 月共计 24 个月的负荷数据作为分析对象，建立贝叶斯网络。数据集中的 5/6 作为学习集来构造贝叶斯网络结构，而剩下的 1/6 作为测试集来检测所构造的贝叶斯网络结构的预测效果。通常认为预测值与实际值的相对误差在 5%以内时，则本次预测准确。贝叶斯网络预测法与多元线性回归预测法的预测效果如表 1 所示。另外，随机对某日的负荷数据进行预测，所以给定该日的 5 个关联参数状态计算负荷预测量属于哪个区间，最后预测负荷与实际负荷进行了对比，对比图如图 4 所示。

表 1 预测结果对比

方法	测试集数	准确预测数	准确率 (%)
多元回归线性法	5 760	4 879	84.7%
贝叶斯网络法	5 760	5 012	87.0%

从表 1 的预测结果对比来看，贝叶斯网络预测法的准确率达到 87%，相对于多元回归线性预测法的 84.7%的准确率来讲有一定程度的提升，证明了所提出的基于贝叶斯网络的短期负荷预测方法的有效性及其一定的优越性。

图 4 为某日 30 min 级预测结果对比。从图

4 中可看出，贝叶斯网络的预测值与实际值较为接近，曲线也较为贴合。多元线性回归法的预测曲线与实际值相比有较大的偏差。

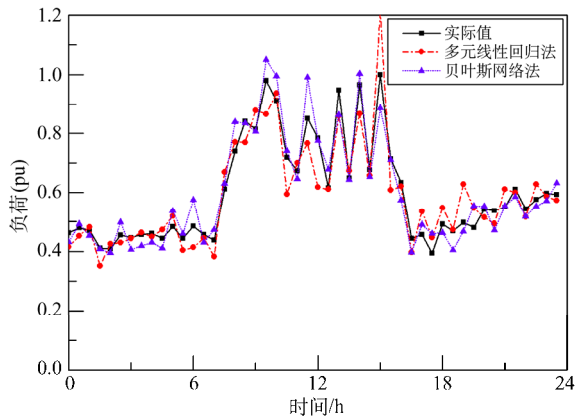


图 4 某日负荷预测效果对比图

式 (12)~式 (14) 为统计中常用的误差指标。误差均值 (Mean Error, ME)，可衡量预测结果是否无偏，即

$$e_{ME} = \frac{1}{n} \sum_{i=1}^n \frac{e_i}{L_i}$$
 (12)

式中， L_i 为实际负荷值。

绝对值平均误差 (Mean Absolute Error, MAE)，可对预测误差的平均幅值进行评价，即

$$e_{MAE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{e_i}{L_i} \right|$$
 (13)

平均相对误差 (Mean Relative Error, MRE)，可用于不同预测算法之间的分析比较，应用也较为广泛。

$$e_{MRE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{e_i}{L_i} \right|$$
 (14)

表 2 为预测误差对比表。从表 2 中可以看出，基于贝叶斯网络的短期负荷预测方法在 3 个统计指标中均优于传统线性回归预测方法，误差能够保持在合理的范围，比传统方法准确度高。

表 2 预测误差对比表

方法	EME	EMAE	EMRE
多元回归线性法	-0.013 4	0.086 1	0.083 4
贝叶斯网络法	0.003 4	0.067 3	0.068 6

4 结束语

将贝叶斯网络应用于电力系统短期负荷预测，给出了贝叶斯网络的数学模型和简要构建方法，归纳了常见的影响负荷的因素，并根据关联因素（参数）推导网络中相应的概率，从而能够确定负荷预测处于何种状态，实现负荷预测。算例分析表明，基于贝叶斯网络的短期负荷预测方法是有效的，且较之多元回归线性预测法具有一定的优越性。贝叶斯网络具有很强的不确定推理能力，同时具备数据挖掘的功能，通过历史数据的分析得出统计规律，符合大数据分析的趋势，该方法可为负荷预测提供一种新的预测思路。

参考文献

[1] 廖旋焕, 胡智宏, 马莹莹. 电力系统短期负荷预测方法综述[J]. 电力系统保护与控制, 2011, 39 (1): 147-152.

[2] 张翠芝, 智明. 泰州电网负荷特性分析及符合预测[J]. 江苏电机工程, 2011, 30 (4): 45-47.

[3] 张林, 罗晓初, 徐瑞林, 等. 基于时间序列的电力负荷预测新算法研究[J]. 电网技术, 2006, 30 : 595-599.

[4] 彭鹏, 彭佳红. 基于多元线性回归模型的电力负荷预测研究[J]. 中国安全生产科学技术, 2011, 7 (9): 158-161.

[5] 朱斌, 刘晓军. 应用人工神经网络算法进行短期负荷预测[J]. 江苏电机工程, 2006, 25 (1): 57-59.

[6] 高山, 张凌浩, 李军红, 等. 节假日短期负荷预测的一种实用算法[J]. 江苏电机工程, 2002, 21 (2): 19-21.

[7] 张素香, 赵丙镇, 王风雨, 等. 海量数据下的电力负荷短期预测[J]. 中国电机工程学报, 2015, 35 (1): 37-42.

[8] 黄影平. 贝叶斯网络发展及其应用综述[J]. 北京理工大学学报, 2013, 33 (12): 1211-1219.

[9] 郑翠芳. 基于贝叶斯网络的软件缺陷预测技术研究与应用[D]. 绵阳 :中国工程物理研究院, 2006.

[10] 刘学艺, 刘祥官, 王文慧. 贝叶斯网络在高炉铁水硅含量预测中的应用[J]. 钢铁, 2005, 40 (3): 17-20.

[11] 郑长江, 葛升阳, 郑树康. 基于贝叶斯网络的交通事件持续时间预测[J]. 华东交通大学学报, 2014, 31 (5): 50-55.

[12] 徐计, 张桂芸. 基于贝叶斯网络的一种牛奶产量预测研究[J]. 计算机工程与科学, 2008, 30 (10): 15-18.

[13] 韩焱红, 矫梅燕, 陈静, 等. 基于贝叶斯理论的集合降水概率预报方法研究[J]. 气象, 2013, 39 (1): 1-10.

[14] 朱明敏. 贝叶斯网络结构学习与推理研究[D]. 西安: 西安电子科技大学, 2013.

[15] 董雷, 周文萍, 张沛, 等. 基于动态贝叶斯网络的光伏发电短期概率预测[J]. 中国电机工程学报, 2013, 33 (2): 38-45.

作者简介

周玉 (1982-), 男, 工程师, 主要研究方向为计量管理、用电信息采集。

(收稿日期: 2015-12-10) EA
(编辑: 崔岩)

(上接第 129 页)

[4] 成大先. 机械设计手册[M]. 北京: 机械工业出版社,

[5] 田雨, 王成林. 有源交变磁场导向定位方法及实验研究[J]. 电子测量技术. 2011, 34 (10): 4-7.

[6] JJF1245.3—2010 安装式电能表型式评价大纲特殊要求 静止式有功电能表 (0.2S、0.5S、1 和 2 级) [S].

[7] Colonel Wm. T Mclyman. 变压器与互感器设计手册[M]. 4 版. 北京: 中国电力出版社, 2014.

[8] Slawomir , Tumansk . Handbook of Magnetic

Measurements[M]. 北京 :机械工业出版社, 2011.

[9] JJG596—2012 电子式交流电能表检定规程[S].

[10] 李宝树. 电磁测量技术[M]. 北京: 中国电力出版社, 2007.

作者简介

熊浩 (1962-), 男, 高级工程师, 主要研究方向为电测计量。

(收稿日期: 2015-12-10) EA
(编辑: 崔岩)