基于离散Fréchet距离的供热效果评价模型

孙志伟**1**，冯海波**1**，王富全**2**，单渊博**1**，董亮亮**1**

（1. 天津科技大学计算机科学与信息工程学院，天津 300457；2. 宏源热力，天津 300457）

**摘　要**：燃煤燃气供热是我国北方主要的供暖方式，而对供热锅炉运行参数的分析将有利于提高供热效果。通过分析供热锅炉系统在运行过程中的数据特点，结合实时采集的气象数据，建立了一个基于离散Fréchet距离以及延时度量和温差度量的供热效果评价模型。结果表明此评价模型对于供热锅炉系统的输出参数（出水温度）及参考曲线的相似度有较好的评价结果，可以为锅炉管控人员提供决策支持和分析依据。

**关键词：**曲线相似度；离散Fréchet距离；延时度量

**中图分类号：              文献标志码：                   文章编号：1672-6510（0000）00-0000-00**

**Evaluation Model of Heat Supply Effect based on Discrete Fréchet Distance**

Sun Zhiwei1，Feng Haibo1，Wang Fuquan2, Shan Yuanbo1，Dong Liangliang1

(1. College of Computer Science and Information Engineering, Tianjin University of Science & Technology, Tianjin 300457, China)

**Abstract:** Heating boiler system is a nonlinear complex system, there are many problems in the practical application. Through the analysis of the boiler system in the running process of the characteristics of the data, combined with real-time acquisition of the meteorological conditions, established a evaluation model based on fréChet distance and increased delay metric to optimize fréChet for lack of time series data processing. Finally, and cluster analysis were combined to verify. The results show that the evaluation model for boiler system output parameters (water temperature) have better evaluation results can provide decision support and basis for the relevant personnel of the boiler control.

**Key words:** boiler control；data mining；Fréchet；cluster analysis；

我国是一个能源消费大国，同时也在能源匮乏的国家之列，北方地区的冬季主要以燃煤燃气供热为主，供暖用能占北方城镇住宅用能总量的60%，具有巨大的节能潜力[1]。而供热锅炉的调控直接影响了能源转换的效率，所以如何根据已有的供热锅炉历史数据建立评价模型或者进行数据分析，辅助锅炉管控人员决策，使锅炉系统提供较好的供热效果，进而减少过度的能源浪费，最终达到节能减排的目的，成为现在一个亟待解决的问题。

诸多专家学者已经对锅炉运行参数进行数据分析，利用数据挖掘等技术展开了一系列工作。孙群丽等对锅炉运行数据进行关联规则挖掘，提供了几组在不同负荷及外部条件下的最优运行方式与参数控制[2]；郑斌祥等通过建立宝钢能源数据仓库对时序数据挖掘的离群挖掘、相似性挖掘、规则挖掘和趋势挖掘等算法进行了系统深入的研究[3]；路海昌等通过对时间序列进行相空间重构，建立了基于支持向量回归的时序数据预测模型，从而实现对锅炉输出参数的预测[4]。

这些研究主要以传统燃煤锅炉为分析对象，而在“十二五”热电联产“煤改气”的过程中处理燃气锅炉系统时表现不足[12]。而将锅炉系统的出水温度与气象条件结合起来，加上延时和温差的处理将极大的提高燃气锅炉系统供热效果的评价结果。综上所述，利用供热锅炉系统的历史数据并结合实时气象条件，加入延时及温差度量共同构建评价模型，可以更有效的为供热锅炉供热效果的判断提供理论依据。

本文通过建立锅炉系统出水温度的参考曲线，利用离散Fréchet距离分析与实际出水温度的相似度，并考虑出水温度的延时及温差度量，建立供热锅炉供热效果的评价模型。通过实验分析，得出了多种控制模式的差异化对比结果，为锅炉系统的相关管控人员提供了决策依据。

# 评价模型的构建

## 1.1 供热参考曲线

锅炉系统管控人员根据多年的运行经验，会制定一套标准作为出水温度的参考，通常为室外温度的线性相关曲线，我们称之为参考曲线。供热锅炉系统的出水温度是依据室外温度等气象因素综合考虑的，为了更好的判断出水温度的合理性，本文将定义一条参考曲线，将室外温度、风速、日照、时间段偏移作为曲线的影响因子，并依据供热公司多年的运行经验采用不同的偏移量得出运行的供热参考曲线:



其各参数解释如下：

t: time，一天当中的某个时刻

Tt: Temperature，t时刻的室外温度

Wt: Wind speed, t时刻的风速

St: Sunshine, t时刻的日照

Pt: Plan，时间段偏移量，依据人们日常活动调整的不同时间段的温度偏移量，如白天上班，夜间休息

a, b, c, d分别为四种条件因素的偏移值，结合供热地区，供热用户特点综合分析设定。

此参考曲线主要有室外温度、风速、日照、时间段偏移四个因素。室外温度、风速以及日照分别是供热区域的气象条件，将直接影响供热用户的采暖需求。而时间段偏移是供热公司依据用户的生活习惯、作息规律进行适度的调控，减少不必要的能源浪费。根据以上分析，综合室外温度、风速、日照以及供热计划四个因素来决定参考曲线的度量。此外，供热锅炉房的出水在流经换热站再抵达采暖用户时会产生一定的传播延时，我们还需要对参考曲线的时刻数据进行归零处理。

## 1.2 曲线相似度比较

出水温度的合理性是依靠室外温度等气象条件来判断的，根据历史数据描绘出两条曲线，分别为上文的参考曲线，以及此时间段内的锅炉系统出水温度曲线。出水温度是根据气象条件实时变化的，且两者走势越符合则表示出水温度越合理，通过判断这两条曲线的相似程度可以来确定供热系统出水温度是否符合期望。此外，根据供热特点，还需要考虑出水温度的及时性，即两条曲线的延时；考虑出水温度是否达到预期及节能效果，即两条曲线的温差。

目前 Hausdorff 距离作为距离的测度被广泛应用于判断两个点集间的相似性，但Hausdorff并未考虑曲线的时序性，而时序在锅炉系统的数据中又是非常重要的参考因素，所以采用Fréchet距离来研究两个曲线相似性[5]。Fréchet 距离由 M. Fréchet提出，描述了两质点分别沿着2条给定曲线以任意速度单向运动时，二者之间的最短距离[9]。Fréchet 距离的直观解释是：给定距离空间的2条曲线A和B，一个人牵着一条狗，分别沿着曲线A和B以任意的自由速度从起点移动到终点，但不得后退，那么 Fréchet 距离就是人与狗之间最短拴狗绳的长度，因此又称狗绳距离。

Axel Mosig和Michael Clausen曾将Fréchet 距离与变换群的交叉子集结合，应用到判别两条曲线的相似性上[6]，曹凯等引入Frechet距离进行云规则推理，设计了一种智能地图匹配算法[11]。

Eiter 和 Mannila在连续 Fréchet 距离的基础上提出了离散 Fréchet 距离的定义 , 而朱洁等考虑了离散Frechet距离的关键特征峰值点，减少了算法的复杂度并将其运用到了手写签名验证上[8]，收到了一定的效果。

提出几种情况的图，否则怎么理解你在干什么。

## 1.4离散Fréchet距离

Eiter和Mannila于1994年在连续Fréchet距离的基础上提出了离散Fréchet距离[10]，本文将离散Fréchet距离作为两条曲线的度量来研究供热锅炉数据，其定义如下：

（1） 给定1个有n个峰值点的多边形链P=<*P1,P2,P3,...,Pn*>，1个沿着P的*k*步，分割P的峰值点成为*k*个不相交的非空子集{Pi}*i=1,…,k*，使得Pi=<Pni-1+1,…,Pni>和1=n0<n1<…nk=n

（2） 给定2个多边形链A=<a1,…,am>，B=<b1,…,bn>，1个沿着A和B的组合步是1个沿着A的*k*步{*Ai*}i=1,…,k和1个沿着B的*k*步{Bi}i=1,…,k组成，使得对于1≤*i*≤*k*，且*Ai,Bi*中有1个恰好包含1个峰值点。

（3） 1个沿着链A和B的组合步*W*={(*Ai,Bi*)}的花费(cost)就是



其中dist()为a,b间的欧氏距离，则链A和B间的离散Fréchet距离就是



## 1.5延时及温差

由于离散Frechet距离可以计算出两条曲线的相似性，可以为温度走势提供依据，但供热锅炉对输入参数产生的反馈需要经历更多的时间延迟，反馈也未必能达到预期值，因此时间参数和温度参数作为参照变量具有相对较高的重要性。本文增加延时度量及温差度量，与离散Fréchet距离共同构成我们评判供热效果的复合度量。

### 1.5.1 延时

两条曲线的延时可以定义为n个同一维度上特征点间的时间差均值。而特征点的确定会对度量结果产生较大影响，考虑到供热锅炉数据的特点，在同一维度上很难找到成对的特征点。这里我们将两条曲线的峰值点作为我们的特征点对，分别计算出两条曲线的极大值，取其时间的差值作为延时，然后计算次大值的延时，依次计算出一段时间内的延时集合，最后将这些延时的均值作为延时度量：



T: 延时度量

i: 特征点的个数

P: 参考曲线

Q: 实际出水温度

### 1.5.2 温差

温差主要表现在预期温度较高时是否能满足采暖需求，温度较低时是否发生能源浪费。这里将分两方面考虑，曲线的极大值差和极小值差。

两条曲线极大值处的温差，能够判断供热效果，出水温度是否能按需达到采暖需求，为供热用户提供足够的热量。参考曲线温度较高时通常意味着用户在家或者气象条件不能提供较多自然热量，所以需要锅炉系统高负荷运转提供充足热量，也可以一定程度上减少客诉。



*T*max: 极大值差

i: 特征点的个数

P: 参考曲线

Q: 实际出水温度

两条曲线极小值处的温差，能够判断供热锅炉系统是否节约能源。参考曲线温度较低时通常意味着，用户家中无人，或者室外温度等气象条件能提供较多自然热量，故锅炉系统需要降低运行负荷减少热量，节约能源降低运行成本。



Tmin: 极大值温差

i: 特征点的个数

P: 参考曲线

Q: 实际出水温度

将极大值处的温差和极小值处的温差复合为一个温差度量，将作为评判锅炉系统是否能提供足够热量以及适时节约能源的依据。



## 1.6 复合度量

将曲线的离散Fréchet距离和延时度量以及温差度量合并为一个复合度量来作为我们供热效果的评价模型:



α: 离散Fréchet距离度量的权重

表1 天津地区供热公司历史数据（部分）

β: 延时度量的权重

γ: 温差度量的权重

ε1: 离散Fréchet距离的阈值

ε2: 延时度量的阈值

ε3: 温差度量的阈值

权重和阈值的确定需要结合具体锅炉房站点通过相应数据进行统计和分析得出。

本文在离散Fréchet距离的基础上加入延时度量及温差度量，作为供热锅炉出水温度和参考曲线的评价模型，将为锅炉系统的供热效果提供理论依据。

# 仿真实验

本文利用天津地区供热公司提供的历史数据，对2015年供热季的数据进行了分析，依据采暖需求，供热公司在夜间会将锅炉系统的出水温度维持在较低的水平，所以我们只对每天6:00至20:00的数据进行分析。由于自动采集系统的稳定性等原因会导致缺失值，错误值等离群点的出现，这里我们先进行数据预处理，通常的缺失值补充方式有，均值法，极值法，拉格朗日插值法，基于时序数据的特点，这里我们采用拉格朗日插值法来进行缺失值的补充，以此作为我们实验的初始数据集，部分数据如表1所示：

|  |
| --- |
| 出水温度/°C 回水温度/°C 气温/°C 日照/Lux |
| 65.88 40.87 6.00 522 66.13 40.86 6.00 522 66.25 40.88 5.90 516 66.38 40.88 5.80 519 66.63 40.88 5.70 516 66.75 40.88 5.60 521 66.88 40.88 5.60 526 66.86 40.87 5.60 522 66.76 40.88 5.60 521 66.88 40.88 5.70 516 |

现依据（1）式将部分数据计算出参考曲线，其中某日的A锅炉房及其参考曲线的对比如图1所示。

参考曲线在上午7时达到极大值，为一天最冷的时间，用户在此时间段在家居多，所以需要提供较高热量。随着室外温度、日照等气象条件的提高，出水温度将逐渐降低，而在下午4时开始，供热公司考虑到更多用户将会回到家中，且室外温度和日照逐渐降低，所以调控锅炉出水温度增加偏移量，为用户提供更多热量。

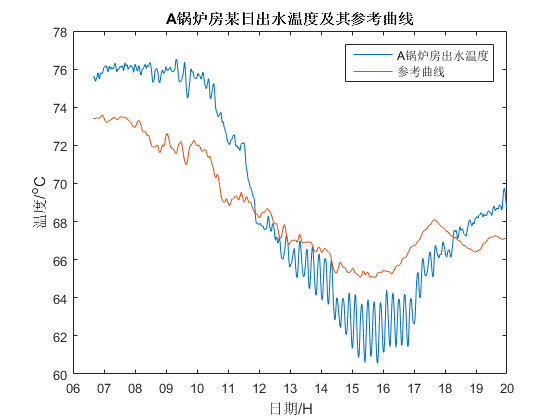


图1 A锅炉房某日出水温度及其参考曲线

然后我们将A站的出水温度曲线与B站的出水温度曲线分别与参考曲线进行相似度对比，依据（2）式、（3）式计算出相应的离散Fréchet距离度量与延时度量，其结果如图2所示：

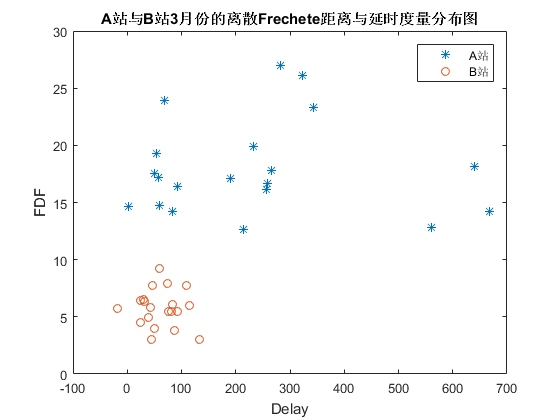


表2 A站与B站3月份耗气量数据表

图2 A站与B站3月份的离散Fréchet距离与延时度量分布图

由此图可知，B站在离散Fréchet距离和延时度量上较A站有更好的表现，在对数据库中的所有数据进行分析，并通过大量的实验验证以及锅炉供热系统相关管控人员历史经验，确定（7）式中的复合度量参数值分别为：

α = 0.6, β = 0.2, γ = 0.2, ε1 = 200, ε2 = 400, ε3 = 10

依据（7）式将上文所述离散Fréchet距离和延时度量以及温差度量复合成为一个评价结果，如图3所示：

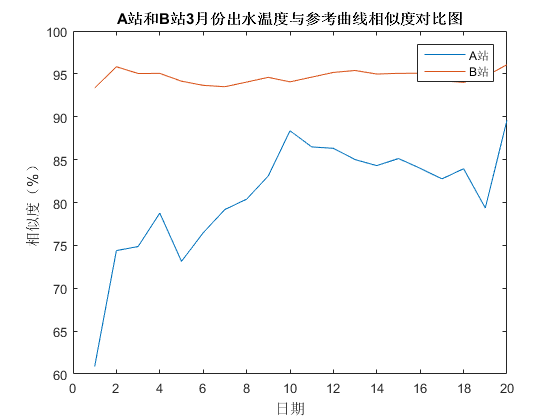


图3 A站与B站3月份出水温度与参考曲线评价结果对比图

此图可以看出B站整体上较A站相比在我们的评价模型上表现更为优秀，取出一天的相应数据作为特征点加以验证，如表2所示

|  |
| --- |
| 锅炉房 总耗气量/m3 供热面积/m2 单位耗气量/°C |
| A 3551056.00 1000.00  B 561134.00 1000.00 |

**DTU网站上供热面积数据不准确**

根据数据可知，A站的单位耗气量大于B站。实际运行中B站采取了自动控制系统来调节出水温度，在考虑气象条件的同时还按照用户的作息规律在相应的时间段设置了温度偏移量，所以B站在与参考曲线的相似度、延时及温差上更为接近，而A站只是按照室外温度进行人工控制，其时效性表现较差。而其实际意义是B站一定程度上减少了资源浪费，节约了能源，这将对锅炉供热系统的相关管控人员提供极大的参考依据，对节约资源浪费具有重大意义。

# 结论

为了保证在供热锅炉系统提供较好的供热效果基础上达到节能减排的目的，本文建立了一个供热效果的评价模型。将离散Fréchet距离和延时度量以及温差度量作为评价因子，分析锅炉系统出水温度曲线和参考曲线的关系，可以很好将不同供热锅炉系统按其供热效果区分出来。利用天津地区某供热公司提供的历史数据，对2015年供热季的数据进行分析，通过本文提出的供热效果评价模型将A站和B站进行了实验，结果表明，通过本文提出复合度量可以将人工控制和自动控制的锅炉房区分出来并判断出较好的供热效果，这为供热锅炉系统的相关人员提供很好的决策依据，对锅炉系统节能减排具有较大意义。此外，锅炉系统的回水温度也是供热效果的评价的因素之一，将出水和回水温度结合起来构建评价模型将会在今后的工作中继续完善。

**参考文献**

[1] 江亿, 彭琛, 胡姗. 中国建筑能耗的分类[J]. 建设科技, 2015(14):22-26.

[2] 张珊. 供热锅炉绩效评价及优化系统的研究[D]. 大连海事大学, 2013.

[3] 郑斌祥. 基于数据仓库的时序数据挖掘研究[D]. 上海交通大学, 2002.

[4] 路昌海, 刘贵松, 张明琤. 基于支持向量回归的锅炉出水温度时间序列预测[J]. 区域供热, 2014(6):18-22.

[5] HELMUT ALT, MICHAEL GODAU. COMPUTING THE FRÉCHET DISTANCE BETWEEN TWO POLYGONAL CURVES[J]. International Journal of Computational Geometry & Applications, 1995, 5(1):75-91.

[6] Axel Mosig, Michael Clausen. Approximately matching polygonal curves with respect to the Fréchet distance ☆[M]// STACS 2001. Springer Berlin Heidelberg, 2010:63-74.

[7] Holm L, Sander C. Mapping the protein universe.[J]. Science, 1996, 273(5275):595-602.

[8] 朱洁, 黄樟灿, 彭晓琳. 基于离散Fréchet距离的判别曲线相似性的算法[J]. 武汉大学学报:理学版, 2009, 55(2):227-232.

[9] Fréchet M M. Sur quelques points du calcul fonctionnel[J]. Rendiconti Del Circolo Matematico Di Palermo, 1906, 22(1):1-72.

[10] Eiter T, Mannila H. Computing discrete Fréchet distance. See Also[J]. See Also, 1994, 64(3):636-637.

[11] 曹凯, 唐进君, 刘汝成. 基于Fréchet距离准则的智能地图匹配算法[J]. 计算机工程与应用, 2007, 43(28):223-226.

[12] 江亿, 唐孝炎, 倪维斗,等. 北京PM2.5与冬季采暖热源的关系及治理措施[J]. 中国能源, 2014, 36(1):7-13.