基于离散Fréchet距离的供热效果评价模型

孙志伟**1**，冯海波**1**，王福全**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）

**摘　要**：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减少能源浪费。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距离的供热效果评价模型，并增加延时度量来优化离散Fréchet距离，从而弥补对供热数据处理能力的不足。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）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；数据挖掘；Fréchet；聚类分析

**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，我国北方地区的冬季主要以燃煤燃气供热为主，且我国供热效率偏低，单位面积供热能耗是同纬度发达国家的3~4倍[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。国内对于锅炉系统的数据分析有很多成果，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]。

这些大多对数据本身分析为主，并未结合气象条件等实际情况，导致评价模型及相关分析不具有普适性，在处理一些复杂情况的锅炉系统时表现不足。而将锅炉系统的出水温度与气象条件结合起来，加上延时的处理将极大的提高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关系，并增加出水温度的延时度量，建立供热锅炉供热效果的评价模型。通过仿真实验，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。

# 锅炉控制系统结构

燃气供热锅炉系统，是通过燃气锅炉加热水循环，输出热水至换热站，经过换热处理向各单位供暖。在燃气供热锅炉控制现场，主要的输出参数有锅炉出水温度、出口总管温度、出水压力、出口总管压力、回水温度、入口总管温度、回水压力、入口总管压力等，而依据室外温度、风速、日照、活动时间等因素确定系统的出水温度，当室外气象条件变化时，控制不同偏移值，出水温度将由自动或者手工方式进行相应调控，以达到用户的供暖需求，因此评判一个供热锅炉系统的优劣的标准可以通过出水温度是否与气象复合因素的变化关系来确定。

# 评价模型的构建

## 2.1 参考曲线定义

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、供热计划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我们的参考曲线，其表达式如下:



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，供热计划，依据日常活动调整的供热变化时间段

a, b, c, d分别为四种条件因素的偏移值，结合地区，供热规模具体给出。

此参考曲线主要有室外温度、风速、日照、供热计划四个因素。室外温度直接影响了系统的出水温度；风可以带走一部分热量，所以按照风力分级对出水温度进行相关调控；由于各地采光条件不尽相同，且在日照强的情况下会增加温感，所以也需按照光照强度对出水温度进行适当调控；用户的作息生活规律是固定的，所以供热计划需要对是否工作日，是否在家的时间段等条件对出水温度进行适度调控。

**延时**

## 2.2 曲线相似度比较

由于出水温度的合理性是依靠室外温度等气象条件来判断的，所以我们根据历史数据描绘出两条曲线，分别是某段时间的气象条件，即上文的参考曲线，另外一条是此时间内的锅炉系统出水温度变化曲线，由于出水温度是根据气象条件实时变化的，且两者振幅与相位越符合则表示出水温度越合理，即通过判断这两条曲线的相似程度来确定供热系统出水温度随参考曲线的应变程度。

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 距离应用到判别两条曲线的相似性[6]，并且Holm L和Sander C还将Fréchet 距离应用到判断蛋白质结构排列的相似性[7]。**（增加其他因素）**

Eiter 和 Mannila在连续 Fréchet 距离的基础上提出了离散 Fréchet 距离的定义 , 而朱洁等人其运用到了手写签名验证上[8]，收到了一定的效果，但离散Fréchet距离过滤了两条曲线的延时条件，所以并不能完全体现供热锅炉系统的曲线的特点，我们将在下文中进行相应的改进形成一个新的复合度量。

## 2.3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距离作为曲线中的关键特征之一来研究曲线的相似性，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



## 2.4离散Fréchet距离及延时的复合度量

由于离散Frechet距离仅能考察峰值点间的距离，这对判断曲线的相似性是远远不够的，并且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，因此时间参数作为一个参照变量具有相对较高的重要性。所以本文增加延时度量，与离散Fréchet距离共同构成我们评判供热效果的复合度量。

两条时序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。而特征点的确立会对度量结果产生较大影响，考虑到供热锅炉数据的特点，实际出水温度往往达不到或者超过预期温度，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，这样有效规避了温度差值的问题。

首先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

将曲线的离散Fréchet距离和延时度量合并为一个复合度量来作为我们出水温度曲线和参考曲线的相似度，构建出水温度的评价模型:



α: 离散Fréchet距离度量的权重

β: 延时度量的权重

ε1: 离散Fréchet距离的阈值

ε2: 延时度量的阈值**（阈值如何确立）**

本文在离散Fréchet距离的基础上加入延时度量，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提取，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极低的水平，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：

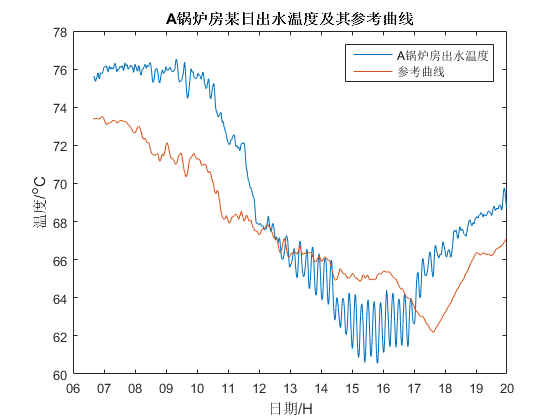


图1 A锅炉房某日出水温度及其参考曲线

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，首先依据（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：

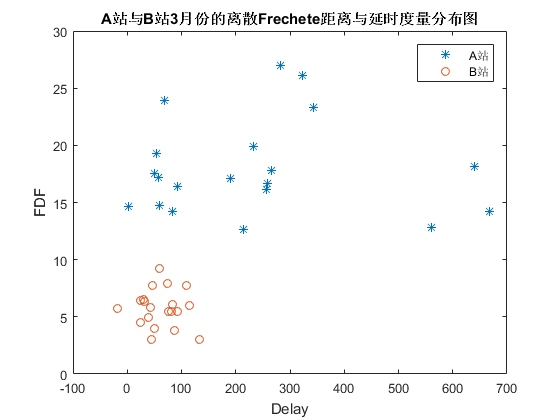


图2 A站与B站3月份的离散Fréchet距离与延时度量分布图

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（4）式中的复合度量参数值分别为：

表1 天津地区供热公司历史数据（部分）

α = 0.8, β = 0.2, ε1 = 200, ε2 = 400

现依据（4）式将上文所述离散Fréchet距离和延时度量复合成为一个度量，如图3所示：（解释图）

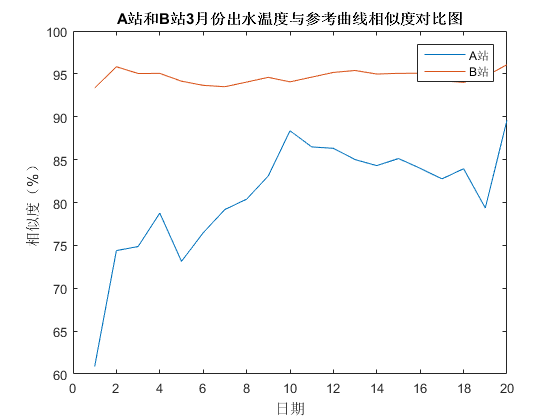


图3 A站与B站3月份出水温度与参考曲线评价结果对比图

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相似，现取出一天的相应数据作为特征点加以验证，如图4所示（耗气量）

**查看具体数据**

图4 A站与B站3月14日出水温度及参考曲线对比图

回溯数据可知，B站采取了自动的智能控制系统来调节出水温度，在考虑室外温度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，而A站只是按照室外温度进行了相应的人工控制，所以在与参考曲线的相似度及延时度量上更为接近，而其实际意义是B站只是根据室外温度调节出水温度，而A站可以根据室外温度、日照等复杂气象条件及时调节出水温度，表现为灵敏度高且反馈及时并能按照人类活动时间段下调出水温度，极大的减少了资源浪费，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经验，对节约资源浪费具有重大意义。

# 结论

为了使供热锅炉系统提供更好的供热效果，本文建立了一个供热效果的评价模型。将离散Fréchet距离和延时度量作为评价因子，判断锅炉系统出水温度曲线和参考曲线的相似性，可以很好将不同的供热锅炉系统按其供热效果区分出来。利用天津地区供热公司提供的历史数据，对2015年3月份的数据进行分析，通过本文提出的供热锅炉供热效果评价模型将A站和B站进行了评价，结果表明，本评价模型通过复合度量可以采用人工控制和自动控制的锅炉房区分出来，这为供热锅炉系统的相关人员提供了很好的决策依据，对相关系统节能减排具有较大意义。由于出水温度受多个复杂的气象条件影响，本文的复合气象条件评价曲线上还有提高的空间，这将在接下来的工作中继续完善。

**参考文献**

[1] 江亿. 我国建筑耗能状况及有效的节能途径[J]. 暖通空调, 2005, 35(5):64-64.

[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.