

基于遗传程序设计的供水管网爆管预测模型

刘俊, 俞国平

(同济大学 环境科学与工程学院, 上海 200092)

摘要: 收集了中国南方某城市2004—2008年的爆管资料, 通过统计分析, 确定了不同管径平均管龄、抗拉强度以及故障强度, 并结合供水管网模拟, 确定了管段的平均绝对压力。通过以管径为分类依据的不同训练样本, 利用遗传程序设计建立了故障强度与管径、平均绝对压力、平均管龄和抗拉强度之间相关关系的爆管预测模型。运用2009年统计爆管数据验证了预测模型的准确性与可行性, 结果表明爆管预测模型能较好地反映实际情况。

关键词: 遗传程序设计; 供水管网; 爆管; 预测模型

中图分类号: TU 991.36

文献标识码: A

Predictive Pipe Failure Models for Water Distribution System Based on Genetic Programming

LIU Jun, YU Guoping

(College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China)

Abstract: The pipe failure data from 2004 to 2008 in a southern city were collected and average pipe age, tensile strength and pipe failure rate were obtained by data analysis, average pipe pressure was obtained by water distribution simulation. With the diameter classification of training samples, predictive pipe failure models which described the pipe failure rate correlations among diameters, average pipe pressure, average pipe age and tensile strength were established on the basis of the genetic programming. The feasibility and effectiveness of the models were verified by the pipe failure data from 2009. The results show that predictive pipe failure models can well describe actual conditions.

Key words: genetic programming; water distribution system; pipe failure; predictive model

爆管是指由于管道的结构性损坏, 管道漏水上升到地面, 必须立即进行维修的情况。供水管道爆管不仅造成大量水资源浪费, 还导致供水经济效益下降, 严重威胁供水安全, 影响人们正常的生产和生活。

许多学者对爆管的因素做了研究, 包括管道属性、外部环境和管理等, 提出了相应的防治对策^[1-4], 定量分析了爆管的众多影响因素^[5], 并借助GIS空间分析技术进行爆管分析^[6-7]。在爆管模型研究方面, 主要包括爆管预测的物理模型和统计模型。物理模型试图分析作用在管道上的荷载, 管道抗荷载的能力, 管道内外所受腐蚀的程度、范围等来预测管道事故^[8-9]。统计模型是以管网运行的爆管数据记录为依据, 利用统计的方法建立爆管事故量化规律^[10]。近几年, 一种新型建模方法, 基于人工智能的数据驱动建模技术受到重视, 通过对数据的学习, 能够获得系统不同变量之间复杂、非线性的相关关系。数据驱动建模技术在爆管预测研究方面已有应用, 并建立了基于人工神经网络和适应性神经-模糊推断系统的爆管模型^[11]。

传统爆管模型需要大量的数据, 如管道特性数据, 准确的和时间足够长的管道维护历史记录等, 但是由于中国供水企业近几年才重视爆管数据的记录与维护, 因此利用现有数据进行爆管分析以及预测的研究较少; 而且, 现有爆管模型一般仅考虑了管径、管龄以及管长等一些基本因素。本文利用南方某市2004—2008年爆管统计资料, 通过遗传程序设计建立了管径、管龄、抗拉强度、管段压力的爆管预测模型, 并以2009年爆管统计资料进行验证。为管网管理以及优化调度提供了依据。

收稿日期: 2009-12-03

基金项目: 国家水体污染控制与治理科技重大专项资助项目(2009ZX07421-005)

第一作者: 刘俊(1982—), 男, 博士生, 主要研究方向为分区供水优化设计与漏损控制。E-mail: sdlj2008@163.com

通讯作者: 俞国平(1945—), 男, 教授, 博士生导师, 主要研究方向为给水排水工程设计与运行最优化。E-mail: gp_yutj@tom.com

1 遗传程序设计

遗传程序设计(Genetic programming, GP)是进化计算方法的一个新分支,由美国斯坦福大学的 JR Koza 教授于 1992 年创立^[12].与遗传算法的基本思想相似,均是随机产生一个适合于给定环境的初始种群,构成群体的个体都有一个适应度值,依据达尔文适者生存原则,采用复制、交叉、变异等处理得到高适应度的个体,产生下一代种群,如此进化下去,给定问题的解或近似解将在某一代出现.与遗传算法不同的是,遗传算法采用二进制编码或实数编码,而遗传程序设计采用数的动态树状结构编码.树的结点由终结点、原始函数与运算符组成.遗传程序设计在符号回归问题上是有用的,而且显示出很强的解题能力.符号回归也称为函数建模,即根据给定的一组自变量和一组函数值,找出拟合函数关系式.

遗传程序设计的基本算子包括复制、交叉和变异.复制依据达尔文的适者生存和自然选择理论对现有群体中的单体进行挑选.根据基于适应度的个体选择方法,从群体中选出一个父代个体,不对其做任何的变化便复制到新一代中去.交叉是从当前群体中,根据适应度值选择两个父个体,两个父个体的不同部件重新组合产生两个子个体.如图 1 所示,两个父代个体代表的表达式分别是 $(+(* 0.23z))$ 和 $(- x 0.76)$ 和 $(* (* z y) (+ y (* 0.31z)))$,随机选择交叉点,如选择第一个父个体的子树 $(* 0.23z)$ 与第二个父代的子树 $(+ y (* 0.31z))$ 进行交叉,得到两个新个体分别是 $(+(+ y (* 0.31z))(- x 0.76))$ 和 $(* (* z y) (* 0.23z))$.

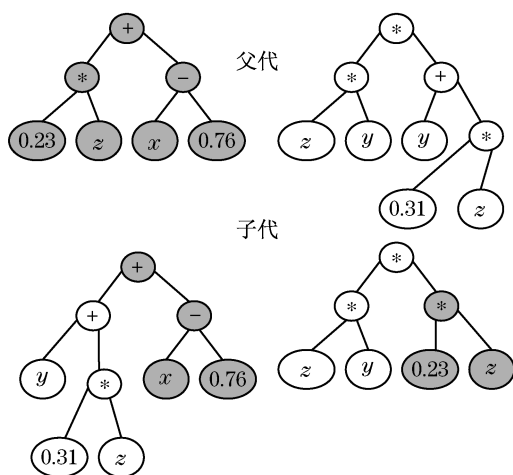


图 1 交叉操作示意图

Fig.1 Sketch map of crossover

变异是由程序随机产生一棵新的子树代替被突变概率选中节点以下的原有子树部分,如图 2 所示.

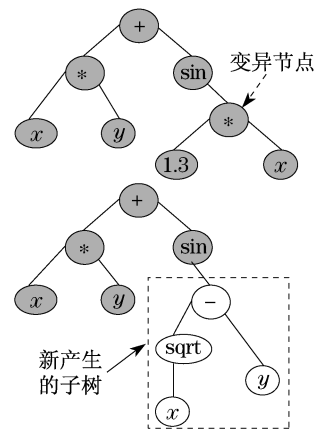


图 2 变异操作示意图

Fig.2 Sketch map of mutation

2 爆管预测模型程序设计方法

2.1 数据收集与整理

收集了中国南方某城市 2004—2008 年 DN600 以上管道爆管统计数据.其供水干管的主要管材包括钢管(50%)、砼管(36%)和球墨铸铁管(11%).

为了便于分析,将所有管段按管径 D 进行分组.计算每组的总管长为 L ,基于管长的加权平均管龄为 A ,基于管网模拟结果的管道流量加权平均绝对压力为 H ,基于管长的加权平均抗拉强度 σ_b ,以及每年每单位管长平均爆管次数即故障强度 λ ^[13],结果见表 1.总管长 L ,加权平均管龄 A ,加权平均绝对压力 H ,加权平均抗拉强度 σ_b 和故障强度 λ 分别由式(1)~(5)计算求出,即

$$L = \sum l_i \quad (1)$$

式中: i 为管段编号; l_i 为管段 i 的长度,Km.

$$A = \frac{\sum A_i \times l_i}{\sum l_i} \quad (2)$$

式中, A_i 为管段 i 的管龄,年.

$$H = \frac{\sum H_i Q_i}{\sum Q_i} \quad (3)$$

式中: H_i 为管段 i 的平均绝对压力,m; Q_i 为管段 i 的流量, $l \cdot s^{-1}$.

$$\sigma_b = \frac{\sum \sigma_{bi} \times l_i}{\sum l_i} \quad (4)$$

式中: σ_{bi} 为管段 i 的抗拉强度, MPa;

$$\lambda = \frac{\sum_{j=i}^n T_j}{n \times \sum l_i} \quad (5)$$

式中: λ 为故障强度, $1 \cdot \text{年} \cdot \text{km}^{-1}$; j 为统计年编号; T_j 为第 j 组统计年爆管次数; n 为统计年总数.

故障强度与许多因素有关, 见表 1, 是多种因素综合的结果, 与任何一个因素没有简单的相关性.

表 1 管道故障强度与相关因素表

Tab.1 Pipe failure rate and related properties

管径 D/mm	平均绝对压力 H/m	平均管龄 $A/\text{年}$	平均抗拉强度 σ_b/MPa	故障强度 $\lambda/1 \cdot (\text{年} \cdot \text{km})^{-1}$
600	39.7	18.4	243	0.22
800	42.0	19.2	226	0.18
1 000	42.4	17.7	269	0.18
1 200	44.9	20.5	199	0.19
1 600	48.1	15.3	375	0.26

2.2 遗传程序设计预测建模

通过遗传程序设计, 求得故障强度与管径、平均绝对压力、抗拉强度和平均管龄之间的相关关系. 基于遗传程序设计的预测分析对表 1 的数据进行预测建模, 为找到最优的爆管预测模型, 采用的模型各参数见表 2.

表 2 遗传程序设计参数设置

Tab.2 Parameters set of genetic programming

参数	取值
终点集	D, H, A, σ_b
因变量	λ
函数集	$+, -, *, /$
种群大小	1 000
进化代数	100
交叉概率	0.60
变异概率	0.01
选择方法	轮盘选择法
终止准则	规定最大允许代数; 误差限内满足所有输入数据组
初始种群产生方法	倾斜对半生成法

根据表 2 所设置的参数, 运行程序. 生成初始种群, 生成含有 1 000 个由 D, H, A 组成的不同表达式. 对第 i 个表达式, 代入第 j 组 D, H, A 的值, 得到计算值 $S(i, j)$, 由式(6)计算原始适应度^[14], 作为该表达式适应度大小的评价标准. 将所有个体按照一

定的概率进行交叉、复制和变异操作, 并记录每个个体对所有输入数组的满足个数, 然后更新种群, 完成第一代进化. 如此进行下去直到满足终止准则, 得到爆管预测模型. 具体流程见图 3.

$$r(i) = \sum_{j=1}^N |S(i, j) - C(j)| \quad (6)$$

式中: $S(i, j)$ 为表达式 i 在第 j 组输入值时的计算值; $C(j)$ 为第 j 组输入值对应的预定输出; N 为输入-输出样本数; $r(i)$ 为表达式 i 在输入所有样本计算得到的绝对误差和.

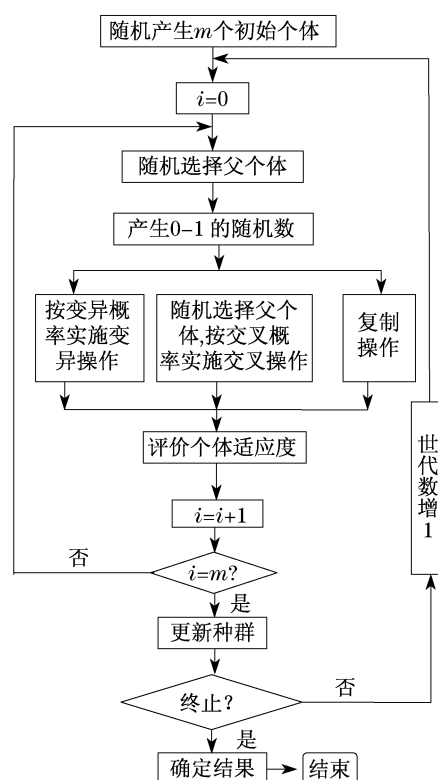


图 3 爆管模型遗传程序设计流程图

Fig.3 Sketch map of genetic programming

3 爆管预测模型模拟分析

3.1 爆管预测模型的建立

对所有样本多次训练以后, 得到若干不同的预测模型. 式(I)、式(II)和式(III)是其中 3 个效果较好的预测模型.

$$\lambda = \frac{\sigma_b}{AH + D - A - \sigma_b} \quad (I)$$

$$\lambda = \frac{D\sigma_b}{D^2 + D\sigma_b + (D + H)(\sigma_b - H)} \quad (II)$$

$$\lambda = \frac{DH}{A(D + \sigma_bH)} \quad (III)$$

3 个预测模型相对误差见图 4. 由图 4 可知, 式 (I) 对 DN600—DN1 000 管道预测效果较好, 相对误差在 5% 以内. 式 II 对 DN600—DN800 管道有较好预测, 但范围较窄, 式 (III) 没有规律. 总体上 3 个预测模型对 DN1200—DN1 600 的管道预测效果不理想. 原因是遗传程序设计是一种全局优化算法, 目标是使所有样本的预测误差最小化. 从图 4 可以看出, DN1 200 以上爆管规律明显与 DN1 000 以下管道不同, 导致总体效果不理想. 因此应以管道为依据对样本分别训练. 管径在 DN600—DN1 000 之间的管道爆管预测模型采用式 (I)

$$\lambda = \frac{\sigma_b}{AH + D - A - \sigma_b} \quad (7)$$

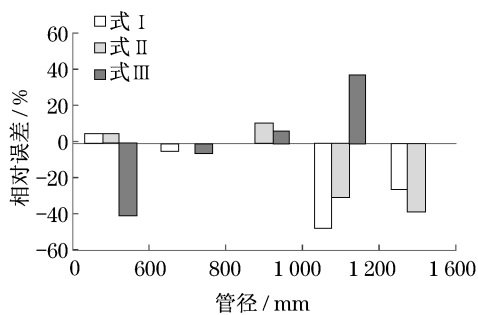


图 4 预测爆管模型相对误差图

Fig. 4 Relative errors of predictive pipe break models

通过对 DN1 200—DN1 600 样本多次训练, 得出式 (8) 预测模型, 其预测值与真实值一致, 即

$$\lambda = \frac{\sigma_b - A + H}{D - H + A} \quad (8)$$

两个不同管径的爆管预测模型模拟效果见图 5. 由图可知 DN600—DN1 000 管道预测相对误差均在 5% 以内, DN1 000—DN1 600 预测误差为零. 总体效果比较理想.

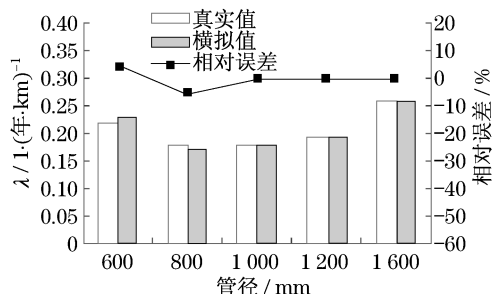


图 5 模拟值与真实值的比较

Fig. 5 Comparison of simulated and measured data

3.2 模型验证

以该市 2009 年户外爆管统计资料验证爆管预

测模型式 (7) 和式 (8) 的准确性与可行性. 效果对比见图 6. DN800—DN1 600 管道预测故障强度与真实故障强度之间相对误差在 10% 以内, 即使相对误差较大的 DN600 管道, 其绝对值也仅相差 4%, 即每 100 km 管道预测爆管次数误差在 4 次以内. 因此, 爆管预测模型能较好的反应实际情况.

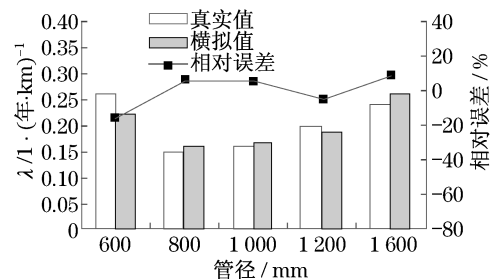


图 6 预测值与真实值的比较

Fig. 6 Comparison of predictive and measured data

4 结论

本文围绕城市供水管网运行中的爆管问题, 运用遗传程序设计建立了多因素爆管预测模型. 主要得出以下结论:

(1) 对大管径输配水管道, 爆管规律有所不同, 训练样本应以管道为划分依据分别训练.

(2) DN600—DN1 000 爆管预测模型为

$$\lambda = \frac{\sigma_b}{AH + D - A - \sigma_b}$$

DN1 200—DN1 600 爆管预测模型为

$$\lambda = \frac{\sigma_b - A + H}{D - H + A}$$

本文通过对爆管资料的统计分析, 结合管网模拟结果, 利用遗传程序设计建立了管径、管龄、抗拉强度及管段压力的爆管预测模型. 分析表明所建爆管预测模型能较好反应实际情况. 可为管网管理以及优化调度提供依据.

参考文献:

- [1] 马乐宁, 刘文君, 徐洪福. 供水管道爆漏事故影响因素实例分析[J]. 给水排水, 2006, 32(9): 86.
MA Lening, LIU Wenjun, XU Hongfu. A case study on factors influencing the burst and leakage of water pipeline[J]. Water & Wastewater Engineering, 2006, 32(9): 86.
- [2] 张玉先, 陈欣, 张硕, 等. 常州市大口径输水钢管爆管原因与对策研究[J]. 给水排水, 2006, 32(7): 89.

(下转第 426 页)