

# 基于贝叶斯理论的城市管网故障检测模型

张方舟<sup>1</sup>, 韩东洋<sup>1</sup>, 周 勃<sup>2</sup>

(1. 东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318;  
2. 大庆采油四场, 黑龙江 大庆 163318)

**摘 要:**随着现在城市信息化建设的加快,城市的正常运转离不开城市地下管网的支持。但是城市地下管网一般规模较大,分布范围广阔,而且故障发生时具有相当的隐蔽性。因此,给城市的发展和社会财富带来严重的隐患。城市管网的故障识别及定位技术一直是近些年来相关企业的研究热门。基于 SCADA 系统检测到的数据,引进统计学概率论中的基本定理—贝叶斯定理来建立管网的在线检测与定位模型,用以帮助对城市管网的运行状态进行分析检测,并对故障进行定位。一定程度上解决了模拟误差、测量误差、测点配置等不确定因素所带来模型检测数据的不确定性问题。

**关键词:**城市管网;在线检测;贝叶斯理论;神经网络

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2015)09-0199-05

doi:10.3969/j.issn.1673-629X.2015.09.042

## Urban Network Malfunction Monitoring Model Based on Bayesian Theorem

ZHANG Fang-zhou<sup>1</sup>, HAN Dong-yang<sup>1</sup>, ZHOU Bo<sup>2</sup>

(1. College of Computer and Information Technology, Northeast Petroleum University,  
Daqing 163318, China;  
2. The Fourth Oil Extraction Plant of Daqing Oilfield Company Limited,  
Daqing 163318, China)

**Abstract:** With the accelerating of informatization in city now, normal operation of a city will not go well without the support of the urban network. However, the scale of urban network is huge, while the network is widely distributed and equipped with concealed problems when fault occurs. Thus, it brings serious hidden troubles to the city's development as well as social fortune. For recent years, the technology of monitoring urban network malfunction and positioning has been a hot research project in related enterprises. Build up a online monitoring and positioning system of network based on data seized by SCADA system and Bayesian theorem which is the basic theorem in statistical probability theory to analyze and detect the running state of urban network as well as positioning malfunction. To some degree the model solves the uncertainty of data detected by model brought by uncertain factors such as simulation error, measurement error or measuring point configuration.

**Key words:** urban network; online monitoring; Bayesian theorem; neural network

## 0 引言

随着我国社会经济的快速发展,城市化进程的加快,同时考虑到城市建设的合理性、美观性因素,因此,将大量的管道、管线改为地下方式组成,城市管网建设也就成了一项必不可少的工作。同时,如何有效地对城市管网的运行状态进行分析检测对城市的正常、稳定发展具有非常重要的意义。

近年来,监控和数据采集系统(SCADA)被应用于

越来越多的城市地下管网中,用以在线监测城市管网中各个控制点的数据信息,如:流量、压力等。这些数据可以帮助管理者更好地监测城市管网的运行状态以及故障信息。因此,根据这些监测数据进行管网的故障实时监测、定位等方面的研究也成了近些年来的一项热门课题。

在国内,对这方面的研究尚在理论阶段。朱东海、梁建文两位学者曾经提出了一种基于 BP 神经网络的

供水管网泄漏检测法,但是这种方法只是对虚拟管网做了研究,且 BP 网络存在着易陷入局部极小点、不易收敛等缺陷,当故障模式较多时,网络结构复杂<sup>[1]</sup>。不过这个理论从一定程度上拓宽了国内管网检测的思路。

## 1 人工神经网络

### 1.1 人工神经网络概述

人工神经网络(Artificial Neural Networks, ANN)是将生物体中神经网络功能利用算法将其模拟出来的一种经验模型。它是一种高度复杂的非线性动力学系统,可以反映出人脑功能中的许多基本特征。ANN 是由许多具有非线性映射能力的神经元组成,通过权系数将神经元和神经元连接<sup>[2]</sup>。它的优势就在于处理非线性映射的能力非常强,在大量的数据中,自动进行归纳总结,并最终获得这些数据的内在规律。

在 ANN 中,权系数和神经元节点是最重要的两个基本概念。其中,节点对应着有向图中的节点,而权系数则代表着节点相互间的连接强度。ANN 具有很强的非线性处理能力和可塑性就在于,连接节点间的权系数是可变的,通过样本训练,学习将权系数不停调整直至达到恰当值,然后输出正确结果。ANN 将训练时的学习成果分化存储于调整后的权系数中。

### 1.2 概率神经网络(PNN)模型

概率神经网络(PNN)是一种常用于模式分类的神经网络。它是一种基于统计原理的神经网络模型,而实质上是通过 Bayes 决策策略中的最小风险准则而发展出来的并行算法。它的整个计算过程是完全向前的,不需要利用 BP 算法对反向误差传播进行计算。因此,训练时间短,不易产生局部最优。而且,它的分类正确率较高,只要有充足的训练样本,无论多么繁杂的分类问题,都可以在贝叶斯准则下得到最优解。

#### 1.2.1 PNN 数学模型

作为模式分类,贝叶斯分类是一种将“期望风险”中最小为优的决定规则,它可以处理大量样本的分类问题。

设训练的模式样本有  $M$  类,整个模式样本空间为  $M \subset R^d, R^d$  是一个  $d$  列的实数空间,  $\vec{x}$  是实属空间的横向量。设第  $i$  维的样本概率密度是  $f_i(\vec{x})$ , 满足

$$\int_M f_i(\vec{x}) d\vec{x} = 1 \quad (i = 1, 2, \dots, M)$$

Parzen 提出,在一般情况下,用于训练网络的模式样本  $P_0, P_1, \dots, P_i$  是已知的,而概率密度未知。将它们视为一种随机变量,其概率密度为  $f_i(\vec{x})$ , 那么在 PNN 中,可以由模式样本来估计条件概率  $p(\mathbf{X} | f_k)$ 。

其估计函数表达式为:

$$p(\mathbf{X} | f_k) = \frac{1}{2\pi^{m/2} \sigma^m n_{p_i \in f_k}} \exp\left[\frac{\mathbf{X}^T \mathbf{W} - 1}{\sigma^2}\right]$$

其中,  $\mathbf{X}$  是待识别样本向量;  $\mathbf{W}$  是模式样本向量;  $m$  是向量维数;  $\sigma$  是平滑因子;  $f_k$  是分类模式;  $n$  是分类模式的数量;  $\mathbf{X}^T$  表示矩阵转置。

这种条件概率的估计方法在模型上就体现为将各个节点上的高斯函数计算并求和。而在每个节点中心,即训练过程中的学习样本,其平滑因子就是该函数的标准偏差  $\sigma$ <sup>[3]</sup>。

#### 1.2.2 基本结构图

根据 Bayes 分类规则与概率密度函数, Specht 提出了 PNN 的层次模型,见图 1。

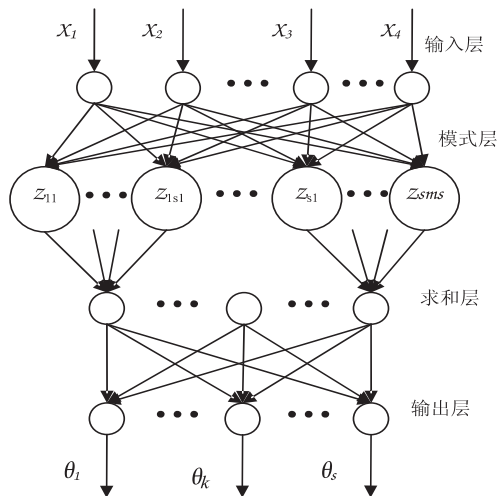


图 1 PNN 结构图

由图可知,在网络训练的过程中,将学习样本的值传入输入层,传输层中的神经元个数等于所有传入的学习样本的矢量维数,再根据经验公式统计法对其高斯函数的平滑因子进行估算。

在模式层里,各种类的训练样本总数与神经元个数相等。将模式层中的各个输入数据的值  $x_i$  分别与权系数  $w_{ki}$  相乘并求和,最后得到  $Z_k$  即输出值。矢量积表示为  $Z_k = \vec{X} \cdot \vec{W}_k$ 。另外,BP 网络中模式层的 Sigmoid 激活函数将采取指数函数  $\exp[(Z_k - 1)/\delta^2]$  来代替。

在求和层中,各个节点分别与其对应模式单元相连接,按照 Parzen 方法,对求和层中的单元求和,并对各个类型的概率进行估算<sup>[4]</sup>。可以看出,贝叶斯决策方法的实现就是由网络的模式层和求和层来完成的。

决策层中,将输入向量根据各类的概率估计,做简单的阈值辨别,在全部决策层神经元中,将具有最大后验概率密度的神经元输出为 1,其他则输出 0。

#### 1.2.3 概率神经网络的优点

(1) 概率神经网络训练容易,收敛速度快,非常适

用于实时处理。

(2)神经网络中模式层采用的是非线性的映射函数,对不同类别的模式样本相互之间的影响都进行了考虑,容错性强。

(3)不易陷入局部最小,在学习训练样本足够的情况下,概率神经网络可以将所有的类别模式都收敛到分类中去。

## 2 基于贝叶斯理论的故障定位机理模型

PNN 的故障检测应用在其他领域,如机械领域中已经颇有成效,因此将其引用于城市管网故障检测中,但是利用该模型对管网进行诊断还有不足:例如故障类型需提前确定;实验时仅能模拟局限的一些故障点及历史的实际故障点<sup>[5]</sup>。

然而,对于泄漏定位来说,故障的类型数量是至关重要的,故障类型越多,泄漏定位的精准度就越高。对于城市地下管网来说,它不同于机械故障诊断领域那样,具有确定的模型故障模式,城市地下管网更加复杂、庞大。若以每个故障点都作为一个故障模式,那么这些故障模式量将会非常庞大。

受实验时实际条件制约,对于故障的特征量描述不够完整,因此存在着很多不足,所以引入新的模型——基于贝叶斯理论的故障定位机理模型。

### 2.1 贝叶斯理论

贝叶斯定理是概率论中的一个基本定理。在将城市管网检测模型的校正时,将管网中的故障信息变化成一定的模型参数(即故障位置和故障数据),将每列模型参数来模拟城市管网中可能的故障事件(如水管泄漏位置和泄漏量)类别,利用贝叶斯方法,将每个故障模式的发生几率进行估算,同时按每个故障模式的相关概率进行排序,发生概率最高的事件被认为是最可能发生的,从而对故障进行检测和定位<sup>[6]</sup>。

### 2.2 建立模型的要点

(1)提高建立模型的准确性。

建立故障定位机理模型,首先需要的是城市管网模型准确、可靠。同时,为了保证模型的精确性以及泛化能力,对模型进行核对校正也是必不可少的。为了提高模型的校核精度,通常在管网模型的校核过程中以反复探测法为主,敏感分析法为辅。这种方法是将模型的参数在已定义的范围内进行调校,并具有相当的广泛实用性和实施性<sup>[7-8]</sup>。校验过程通常是先建立时间框架,再将数据进行比较、分析,如此循环直至比较结果达到允许的精度范围<sup>[9]</sup>。

(2)优化管网的测点。

监测点包括压力监测点、流量监测点、水质监测点等,监测点配置,是地下管网运行管理的一个重要环

节<sup>[10]</sup>。在整个城市管网中,各个监测点的数量要合适,其分布位置也要合理,这样才能正确、精准地反映出整个城市管网系统的运行状态。

管网模型校核过程如图2所示。

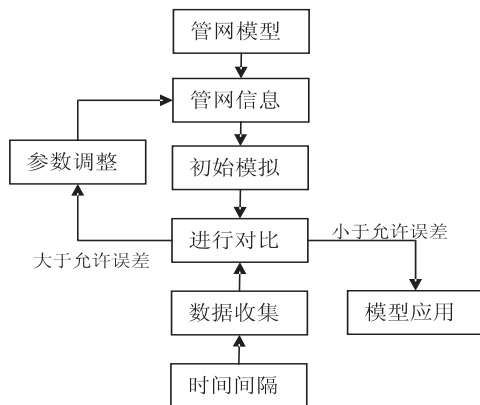


图2 管网模型校核过程

(3)完备的自动监控系统。

目前,城市管网中大部分都已配备了 SCADA 系统<sup>[11]</sup>,管网节点上的监测装置将实时监测管网数据,并远程传送数据到控制中心。随着计算机技术和通信技术的日益发展,这样的监控系统将会越来越完善,它将会成为整个城市管网不可或缺的一部分,将为城市管网安全、协调、高效的运作提供有力保障。

### 2.3 模型的形成

#### 2.3.1 模型原理

利用管网模型中节点处的质量平衡方程、环(包括虚拟环)中的能量守恒方程求解系统工况。需要了解管网的各项参数,如:管长、管径、管网节点用量以及管网拓扑关系等。一般采用牛顿-拉夫森迭代法对模型求解<sup>[12-13]</sup>。

假设引进模拟模型  $M$ , 其中与故障相关的参数(故障数据和故障位置)为  $\theta$ 。那么,只要根据某个可以确定的参数  $\theta$ ,就能够推算出当前管网的行为状态是否正常。

假设,城市管网中装有 SCADA 系统,用其对管网区域中  $K$  个监测点进行监测,并收集分析监测到的  $M$  次数据。那么就可以假定  $\tilde{x}$  是一个大小为  $K \times M$  的矩阵,其矩阵中元素  $\tilde{x}_{ij}$  指的是来自于第  $i$  个监测点检测到的第  $j$  次监测数据。

#### 2.3.2 基于贝叶斯的统计系统辨认法

假设,从模型计算的管网系统某监测点数据为  $x(\theta) = [x_{ij}(\theta)]$ 。那么该节点上通过模型计算的值和测量值之间的偏差为

$$e_{ij}(\theta) = \tilde{x}_{ij} - x_{ij}(\theta) \quad (1)$$

此偏差主要是因为在实际过程中,类似物理仪器测量的误差、模型模拟误差等在模拟过程中一些难以

避免的因素造成的。

在贝叶斯方法中,将这种误差叫做模型预误差 $e_{ij}(\theta)$ 。贝叶斯方法将这种误差看作是一个源自概率误差模型 $\Pi$ 的随机数,模型 $\Pi$ 的参数集为 $\sigma$ ,用参数集 $[\theta, \sigma]$ 表示模型 $M$ 和模型 $\Pi$ 中有意义的参数<sup>[5]</sup>。

统计系统辨认法是想将参数集 $[\theta, \sigma]$ 的值和它们的相关不确定性通过测量值进行修正。利用概率密度函数(PDF)对所有的模型的不确定性进行量化,在没有获取到够多的数据之前,也可以根据项目经验对模型的参数的不确定性进行量化,这就是初始的PDF( $\Pi(\theta, \sigma)$ )。

由 Bayes 理论,根据搜集到的数据将初始 PDF 进行统计计算后,得到新的 PDF 表达式:

$$p(\theta, \sigma | \tilde{x}) = c_1 p(\tilde{x} | \theta, \sigma) \pi(\theta, \sigma)$$

(2)

式(2)是根据足够多的监测数据重新计算了模型的相对确定值。其中 $c_1$ 为一常数,使整体概率趋向于1<sup>[14-15]</sup>。假设 $e_{ij}(\theta)$ 是独立的均值为0、标准偏差为 $\sigma$ 的正态分布,则 $p(\theta, \sigma | \tilde{x})$ 可表达为:

$$p(\tilde{x} | \theta, \sigma) = \frac{1}{\prod_{i=1}^K \prod_{j=1}^M \sqrt{2\pi} \sigma} \exp \left[ - \frac{(x_{ij}(\theta) - \tilde{x}_{ij})^2}{2\sigma^2} \right]$$

(3)

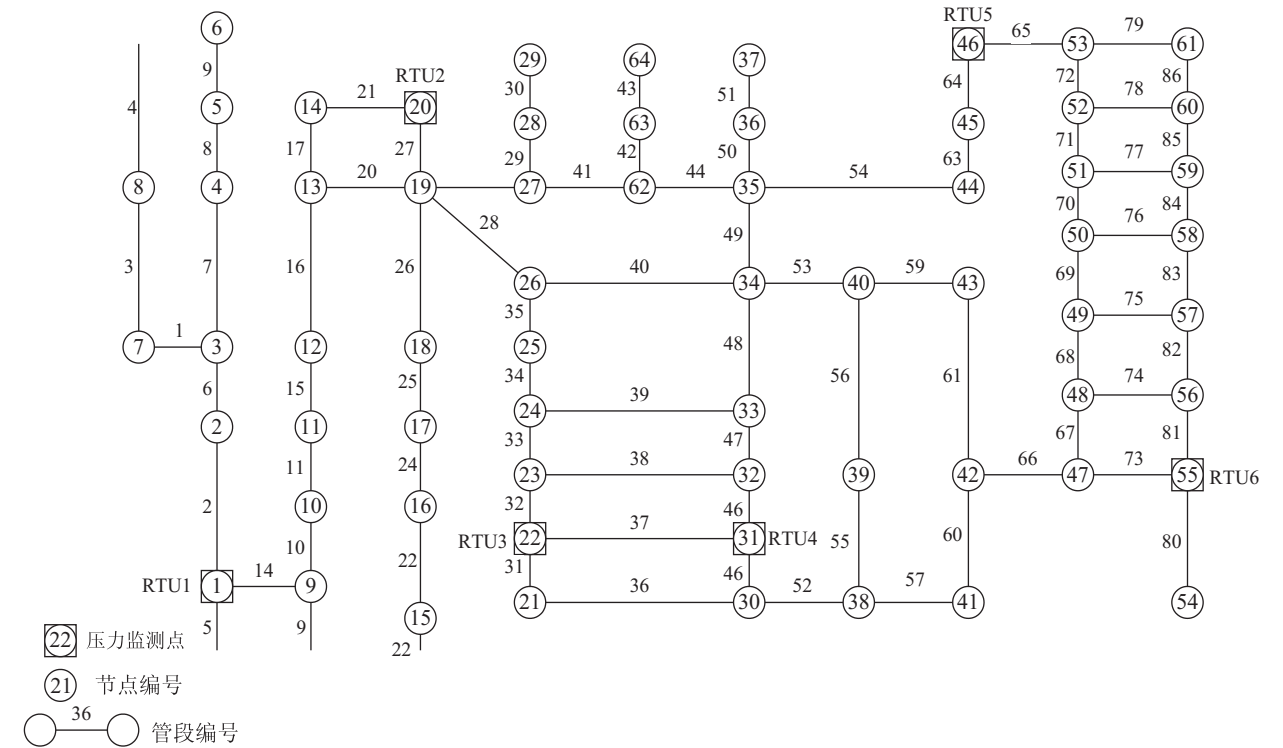


图 3 实例管网示意图(部分)

该管网的某次泄露事故为 49 节点附近 69 管段处,事故发生时水源 1 来水量为 360 m<sup>3</sup>/h,水源 2 来水量为 210 m<sup>3</sup>/h,水源 3 的来水量为 340 m<sup>3</sup>/h。事故发生时,水网中各监测点的水压检测值见表 1。

另外,假定模型的参数集 $[\theta, \sigma]$ 为无信息的先验分布,又设最初的 $\pi(\theta, \sigma)$ 是常数。那么,将式(3)代入式(2)就可得:

$$p(\theta, \sigma | \tilde{x}) = c_2 \frac{1}{(\sqrt{2\pi} \sigma)^{KM}} \exp \left[ - \frac{\sum_{j=1}^M \|x_j(\theta) - \tilde{x}_j\|^2}{2\sigma^2} \right]$$

(4)

式中, $x_j(\theta)$ 为模型计算结果; $\tilde{x}_j$ 为对应检测位置的第 $j$ 次数据检测值; $c_2$ 为常数,用以使概率集成到1。

3 模型在故障检测中的应用

假定一个管网由 $R$ 个管段组成,所有有可能的故障位置数为 $L$ ,那么可以将模型参数集分为两个子集,即 $\theta = (\theta_1, \theta_s)$ 。子集 $\theta_1$ 表示的是有故障的管段,假定故障为管段的中间位置,那么 $\theta_s$ 表示给定的故障处的故障数据。定量化表示在相应的 $L$ 个故障位置的故障程度。

以大庆某小区民用水供水管网为例。该水网共计 64 个节点,86 个管段,周边一共 3 个水源,见图 3。

表 1 泄漏时各监测点数值

	RTU <sub>1</sub>	RTU <sub>2</sub>	RTU <sub>3</sub>	RTU <sub>4</sub>	RTU <sub>5</sub>	RTU <sub>6</sub>
节点	1	20	22	31	46	55
压力值/mH <sub>2</sub> O	24	24	23.6	23.4	18	16



泄漏检测结果见表 2。

表 2 泄漏检测结果

管段 编号	泄漏量 /L · s <sup>-1</sup>	概率	管段 编号	泄漏量 /L · s <sup>-1</sup>	概率
62	14	0.965	61	14	0.947
69	14	0.965	59	14	0.932
68	14	0.965	58	14	0.932
76	14	0.965	72	14	0.932
70	14	0.965	86	14	0.916
71	14	0.947	56	14	0.916
84	14	0.947	53	14	0.909
85	14	0.947	60	14	0.909
83	14	0.947	54	14	0.9
82	14	0.947	63	14	0.9

由表 2 可知,共有五处发生泄漏的概率最大,为 0.965,分别是 62、69、68、76、70 号管段。由图 3 可以看出,这些管段全部都是在真实故障点 69 号管段附近,并与其直接或间接相连。而其他概率的产生都是由于在实际中存在着模拟误差、测量误差、监测误差等诸多不可避免的不确定因素。但是通过模型计算,已经可以将估算的诊断范围缩小到真实的故障区域周围比较精确的范围中了。

多重泄露的检测原理与单一泄露相同,只是计算强度增加。同理,其他类型的地下管网故障,都可以用该模型进行检测。

4 结束语

基于供水管网故障检测和定位存在着诸多不确定因素。引入贝叶斯定理的检测定位模型,已经可以将模型误差、测量误差等存在的实际问题在一定程度上进行修正了。

不过,整个管网的 SCADA 系统和建立的管网模型结构依旧影响着整个模型的准确性。因此,对于已经建立起 SCADA 系统的城市管网,对其数据的准确性需要进一步的确认,并增加一定量的监测点。同时,对管网的模拟模型需要进一步的优化完善,从而提升模型

对于管网检测的适用性、广泛性。

参考文献:

[1] 李 霞,王晓东,赵新华,等. 基于贝叶斯理论的城市供水管网泄漏在线检测与定位[J]. 给水排水,2006,32(12):96-99.

[2] 莫剑冬,徐章遂,米 东. 应用概率神经网络诊断自行火炮发动机的故障[J]. 华北工学院测试技术学报,2000,14(1):7-11.

[3] 郭新宇. 应用概率神经网络诊断自行火炮发动机的故障[J]. 测控技术,2007,26(8):4-5.

[4] 苏建元,王大伟. 复杂电力电子主回路故障诊断研究与仿真[J]. 测试技术学报,2008,22(2):171-174.

[5] 李 霞. 城市供水管网漏损定位及控制研究[D]. 天津:天津大学,2006.

[6] 陈泉博. 燃气管网泄漏工况分析及故障诊断[D]. 哈尔滨:哈尔滨工业大学,2008.

[7] 郑 彤,陈春云. 环境系统数学模型[M]. 北京:化学工业出版社,2003.

[8] de Schaetzen W B F, Walters G A, Savic D A. Optimal sampling design for model calibration using shortest path, genetic and entropy algorithms[J]. Urban Water,2000,2(2):141-152.

[9] Lansey K E, Basnet C. Parameter estimation for water distribution networks[J]. Journal of Water Resources Planning and Management,1991,117(1):126-144.

[10] 黄廷林,丛海兵. 给水管网测压点优化布置的模糊聚类法[J]. 中国给水排水,2001,17(11):50-52.

[11] Cembrano G, Wells G, Quevedo J, et al. Optimal control of a water distribution network in a supervisory control system[J]. Control Engineering Practice,2000,8(10):1177-1188.

[12] Isaacs L T, Mills K G. Linear theory methods for pipe network analysis of hydra[J]. ASCE,1972,98:1191-1201.

[13] 严煦世,赵洪宾. 给水管网系统理论与分析[M]. 北京:中国建筑工业出版社,2003.


[14] 陈 磊,张土乔,吕 谋,等. 遗传算法优化管网神经网络模型[J]. 中国给水排水,2003,19(5):5-7.

[15] 丛海兵,黄廷林. 给水管网状态估计的优化方法[J]. 给水排水,2001,27(9):34-38.

# 基于贝叶斯理论的城市管网故障检测模型

作者：[张方舟](#)，[韩东洋](#)，[周勃](#)，[ZHANG Fang-zhou](#)，[HAN Dong-yang](#)，[ZHOU Bo](#)

作者单位：[张方舟, 韩东洋, ZHANG Fang-zhou, HAN Dong-yang \(东北石油大学 计算机与信息技术学院, 黑龙江 大庆, 163318\)](#)，[周勃, ZHOU Bo \(大庆采油四场, 黑龙江 大庆, 163318\)](#)

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：2015(9)

引用本文格式：[张方舟](#).[韩东洋](#).[周勃](#).[ZHANG Fang-zhou](#).[HAN Dong-yang](#).[ZHOU Bo](#) [基于贝叶斯理论的城市管网故障检测模型](#)[期刊论文]-[计算机技术与发展](#) 2015(9)