

基于置信规则的专家系统用于管道泄漏检测的推理与学习方法

徐东玲^{a,*}, 刘军^b, 杨剑波^a, 刘国平^{c,e,f}, 王进^d, 伊恩·詹金森^d, 任军^d

^a 曼彻斯特大学曼彻斯特商学院, 哈罗德·汉金大楼 502 室

西布斯街, 曼彻斯特 M13 9QH, 英国

^b 英国阿尔斯特大学约旦斯顿校区计算与数学学院

^c 英国格拉摩根大学电气工程学院

^d 英国利物浦约翰摩尔大学工程学院

^e 中国科学院复杂系统智能控制与管理国家重点实验室

^f 中国中南大学网络计算与安全技术湖南省重点实验室

摘要

基于置信规则的专家系统是传统规则系统的扩展, 能够利用具有不确定性的不同类型信息来表示更复杂的因果关系。本文介绍了如何训练基于置信规则的专家系统并将其用于管道泄漏检测。利用专家知识, 通过置信规则库对不同条件下的管道运行进行建模, 然后使用管道运行数据对其进行训练和微调, 并通过测试数据进行验证。所有训练和测试数据均来自一条真实管道并进行了标准化处理。研究表明, 基于置信规则的系统具有灵活性, 能够适应表示复杂的专家系统, 是一种有效的管道泄漏检测新方法。©2005 爱思唯尔有限公司。保留所有权利。

关键词: 置信规则库; 专家系统; 证据推理方法; 泄漏检测; 优化

1. 引言

置信规则库的概念及其推理方法由杨、刘、王、西伊和王 (2005a) 基于证据推理方法 (杨, 2001; 杨和徐, 2002a, 2002b) 提出。在置信规则库中, 规则的每个可能结果都与一个置信度相关联。这样的规则库能够捕捉不同因素之间更复杂和连续的因果关系, 而传统的“如果 - 那么”规则是其特殊情况 (霍奇斯等人, 1999; 帕森, 1996; 孙, 1995)。

应用置信规则库时, 将规则前件的输入转换为前件参考值上的置信分布。然后使用该分布计算规则库中规则的激活权重。随后, 置信规则库中的推理是通过使用证据推理方法对所有激活的规则进行聚合来实现的。

置信规则库中的规则可以通过不同的方式获得, 例如从专家那里提取、通过检查历史数据或通过从训练数据中进行自学习。手动提取规则的过程可能很耗时, 并且规则可能是近似的。因此, 杨、刘、徐、王和王 (2005b) 提出并研究了用于训练置信规则库的最优学习方法。

尽管新方案通过简单的数值示例进行了演示, 但其在处理更实际和复杂问题时的有效性和能力仍需测试。在本文中, 我们将应用该新方案构建一个用于管道泄漏检测的专家系统。

管道泄漏可能会对环境造成不可估量的破坏, 并给管道运营公司带来损失。为了尽量减少这些破坏和损失, 人们开发了许多管道泄漏检测方法和系统类型, 例如基于质量平衡的方法 (Rou-gier, 待发表)、实时传输模型 (Abhulimen & Susu, 2004; Carpenter, Nicholas, & Henrie, 2005)、统计分析方法 (Buchberger & Nadimpalli, 2004) 以及声发射检测方法 (Gao, Brennan, Joseph, Muggleton, & Hunaidi, 2005; Lee, Lee, & Park, 2004)。基于压力校正质量平衡的实时系统是非常受欢迎的系统之一。下面研究的基于置信规则的泄漏检测专家系统也将基于质量平衡原理。

* 通讯作者。电话:+44 0161 2750941; 传真:+44 0161 200 3505。

电子邮箱地址:L.Xu@mbs.ac.uk(徐德龙)。

当管道发生泄漏时,管道内的流量和压力会按照一定模式发生变化。专家能够提供一组规则来区分正常运行和泄漏情况下的模式。传统的基于规则的专家系统或许能够捕捉到用于泄漏检测的这些规则。然而,此类系统可能无法检测到小泄漏,也无法提供关于泄漏大小的准确信息。

本文描述了应用基于置信规则的系统进行管道泄漏检测的过程和结果。在接下来的章节中,将概述构建基于置信规则的专家系统的概念和方法,随后详细描述如何使用置信规则库开发管道泄漏检测专家系统。该系统使用实际管道数据进行训练和测试,为保护管道所有者的利益对数据进行了缩放处理。最后,本文讨论了新系统的优势以及进一步开发的需求。

2. 基于置信规则的专家系统概述

基于邓普斯特 - 谢弗证据理论 (Shafer, 1976)、决策理论和模糊集理论, Yang 等人提出了一种新方法,用于使用置信结构构建混合规则库,并使用证据推理方法在基于规则的系统中进行推理 (Yang, 2001; Yang & Sen, 1994; Yang & Singh, 1994; Yang & Xu, 2002a, 2002b)。该方法被称为基于证据推理方法的通用规则库推理方法 - RIMER (Yang 等人, 2005a)。

2.1. 置信规则库

在 RIMER 方法 (Yang 等人, 2005a) 中,一条置信 IF - THEN 规则,例如第 k 条规则 R_k , 表示如下:

$$\text{IF } (X_1 \text{ is } A_1^k) \wedge (X_2 \text{ is } A_2^k) \wedge \cdots \wedge (X_{T_k} \text{ is } A_{T_k}^k) \\ \text{THEN } \{(D_1, \beta_{1k}), (D_2, \beta_{2k}), \dots, (D_N, \beta_{Nk})\}, \left(\sum_{i=1}^N \beta_{ik} \leq 1 \right),$$

规则权重为 $\theta_k (k = 1, \dots, L)$

属性权重为 $\delta_1, \delta_2, \dots, \delta_{T_k}$ 。(1)

其中 $A_i^k (i = 1, \dots, T_k)$ 是第 i 个前件属性的参考值, T_k 是第 k 条规则中使用的前件属性的数量。如果 $(X_1, X_2, \dots, X_{T_k}) = (A_1^k, A_2^k, \dots, A_{T_k}^k)$ 。 L 是规则库中所有规则的数量, $\beta_{ik} (i = 1, \dots, N)$ 是在 D_i 被认为是结果时的置信度。如果 $\sum_{i=1}^N \beta_{ik} = 1$, 则称第 k 条规则是完整的; 否则, 它是不完整的。请注意, $\sum_{i=1}^N \beta_{ik} = 0$ 表示在第 k 条规则中, 给定输入时对输出的完全无知。规则 (1) 也被称为置信规则。进一步假设 T 是规则库中使用的前件属性的总数。

设

$$X = (X_1, X_2, \dots, X_{T_k}), A^k = (A_1^k, A_2^k, \dots, A_{T_k}^k), D = (D_1, D_2, \dots, D_N)$$

$$\beta^k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{Nk}), \text{ and } \delta = (\delta_1, \delta_2, \dots, \delta_T)$$

X 被称为第 k 条规则的输入向量, A^k 是一个数据包前件, $A_i^k (i = 1, 2, \dots, T_k)$ 是数据包前件的第 i 个参考值, A^k, D 是结果向量, β^k 是置信度向量, δ 是规则库中所有 T 个前件属性的属性权重。

不难看出传统 IF - THEN 规则和置信 IF - THEN 规则之间的区别。在传统规则中, 结果要么 100% 为真, 要么 100% 为假。这样的规则库在表示现实世界中的知识时能力有限。例如, 它无法捕捉前件和结果之间的连续因果关系。置信规则库中的置信结构在表示不同结构和复杂度的知识方面提供了更好的灵活性, 例如前件和结果之间的连续和不确定关系。

2.2. 使用证据推理 (ER) 方法进行规则推理

给定系统的一个输入 $U = (U_i, i = 1, \dots, T)$ ，如何使用规则库进行推理并生成输出？如前所述， T 是规则库中前件的总数， $U_i (i = 1, \dots, T)$ 是第 i 个属性，它可以是以下类型之一 (Yang 等人, 2005a): 连续型、离散型、符号型和有序符号型。

在推理过程开始之前，需要确定输入与规则前件中每个参考值的匹配度，以便为每条规则生成一个激活权重。这相当于使用置信度将输入转换为参考值上的分布，并且可以使用不同的技术来实现，例如基于规则或效用的等价转换技术 (Yang, 2001)。

使用上述符号，第 k 条规则的激活权重 w_k 计算如下 (Yang 等人, 2005a):

$$w_k = \frac{\theta_k \times \prod_{i=1}^{T_k} (\alpha_{ik})^{\bar{\delta}_i}}{\sum_{j=1}^L \left[\theta_j \times \prod_{l=1}^{T_k} (\alpha_{lj})^{\bar{\delta}_l} \right]}, \quad (2)$$

其中 $\bar{\delta}_i = \frac{\delta_i}{\max_{i=1, \dots, T_k} \{\delta_i\}}$, $\alpha_{ik} (i = 1, \dots, T_k)$ 是输入 $X_i \in U$ 与第 i 条规则中数据包前件 A^k 的第 i 个参考值 A_i^k 的个体匹配度， $\alpha_{ik} \geq 0$ 和 $\sum_{i=1}^{T_k} \alpha_{ik} \leq 1$. $\alpha_k = \prod_{i=1}^{T_k} (\alpha_{ik})^{\bar{\delta}_i}$ 被称为组合匹配度。

在确定了规则库中每条规则的激活权重后，可以直接应用证据推理 (ER) 方法 (Yang, 2001; Yang & Xu, 2002a) 来组合这些规则并得出最终结论。假设组合的结果如下:

$$O(U) = \{(D_j, \beta_j), j = 1, \dots, N\} \quad (3)$$

式 (3) 所表达的结果表明，如果输入为 $U = (U_i, i = 1, \dots, T)$ ，那么结果为 D_1 的程度为 β_1 , D_2 为 β_2, \dots 的程度为 β_2, \dots ，为 D_N 的程度为 β_N 。利用证据推理算法的解析形式 (Wang, Yang, & Xu, 待发表)，可以得出 D_j 中的组合置信度 β_j 如下:

$$\beta_j = \frac{\mu \times \left[\prod_{k=1}^L \left(w_k \beta_{jk} + 1 - w_k \sum_{j=1}^N \beta_{jk} \right) - \prod_{k=1}^L \left(1 - w_k \sum_{j=1}^N \beta_{jk} \right) \right]}{1 - \mu \times \left[\prod_{k=1}^L (1 - w_k) \right]}, \quad j = 1, \dots, N \quad (4)$$

其中 $\mu = \left[\sum_{j=1}^N \prod_{k=1}^L \left(w_k \beta_{jk} + 1 - w_k \sum_{j=1}^N \beta_{jk} \right) - (N-1) \times \prod_{k=1}^L \left(1 - w_k \sum_{j=1}^N \beta_{jk} \right) \right]^{-1}$ ，且 w_k 如式 (2) 所示。

2.3. RIMER 中训练置信规则库的最优学习方法

尽管可以通过从专家那里提取知识来建立置信规则库，但如果通过从可用的历史数据中学习对规则进行微调，系统的性能可以得到提升。规则库的可调参数包括置信度 ($\beta_{1k}, \beta_{2k}, \dots, \beta_{Nk}$ ，规则权重 θ_k (对应 $k = 1, 2, \dots, L$) 以及属性权重 ($\delta_1, \delta_2, \dots, \delta_T$) (Yang 等人, 2005b)。

图 1 描绘了训练置信规则库的过程，其中 U 是给定的输入， \hat{O} 是通过仪器测量或专家评估得到的相应观测输出， O 是由置信规则库系统生成的模拟输出， $\xi(P)$ 是 \hat{O} 和 O 之间的差异，以及

$$P = (\beta_{ik}, \theta_k, \delta_j; i = 1, \dots, N; k = 1, \dots, L; j = 1, \dots, T) \quad (5)$$

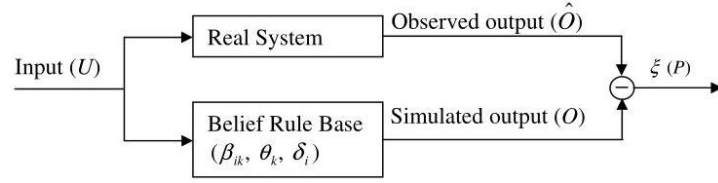


图 1. 最优学习过程示意图。

是可调参数。训练的目标是通过调整参数 P 来最小化差异 $\xi(P)$ 。即使是专家也很难手动实现这一目标，不过有可用的计算机算法来解决这个问题。Yang 等人 (2005b) 更详细地讨论了如何针对不同类型的输出构建问题以及应用何种算法来解决这些问题。

3. 基于置信规则的管道泄漏检测专家系统

3.1. 问题描述

在本节中，我们将考虑一条长度超过 100 km 的管道，其入口和出口处安装有质量流量计，入口、出口以及沿线 8 个中间点安装有压力计。每隔 10 s 从这些仪表收集一次数据。该管道大多在无泄漏 (正常) 状态下运行。然而，在泄漏试验期间，管道中制造了一系列泄漏。每次泄漏持续数小时，并且泄漏大小通过阀门控制。图 2 分别显示了在一次泄漏试验约 5 个半小时内收集的入口和出口流量及压力读数 (f_0, f_1, p_0 和 p_9)，泄漏期间通过入口和出口流量读数的巨大差异清晰标记出来。

我们将使用这些数据来训练和验证一个基于置信规则的专家系统，以检测这些泄漏并估计泄漏大小，同时不产生误报。

3.2. 规则库的前件和后件属性

在正常运行时，当入口流量大于 (或小于) 出口流量时，管道内的压力会升高 (或降低)，因为管道内的总物质量在增加 (或减少)。然而，如果这种模式被打破，例如，入口流量大于出口流量，但管道内的压力仍然降低，那么很可能管道存在泄漏。因此，入口流量与出口流量之差 (用 FlowDiff 表示) 以及管道平均压力随时间的变化 (用 PressureDiff 表示) 是检测管道是否泄漏的两个非常重要的因素。它们是规则库的两个前件属性，计算方法如下：

$$U_1 = \text{FlowDiff}(t) = f_1(t) - f_0(t) \quad (6)$$

$$U_2 = \text{PressureDiff}(t)$$

$$= [p_0(t) + \dots + p_9(t)] / 10 - [p_0(t-1) + \dots + p_9(t-1)] / 10$$

(7) 其中 $f_1(t), f_0(t), p_0(t), \dots$ 和 $p_9(t)$ 是在时间 t 收集的仪表读数。后件属性是泄漏率，用 LeakSize 表示。在泄漏试验期间控制了 LeakSize 的值，因此其值或多或少 (尽管并非完全) 是已知的。表 1 列出了一些前件属性值和相应的后件属性值。

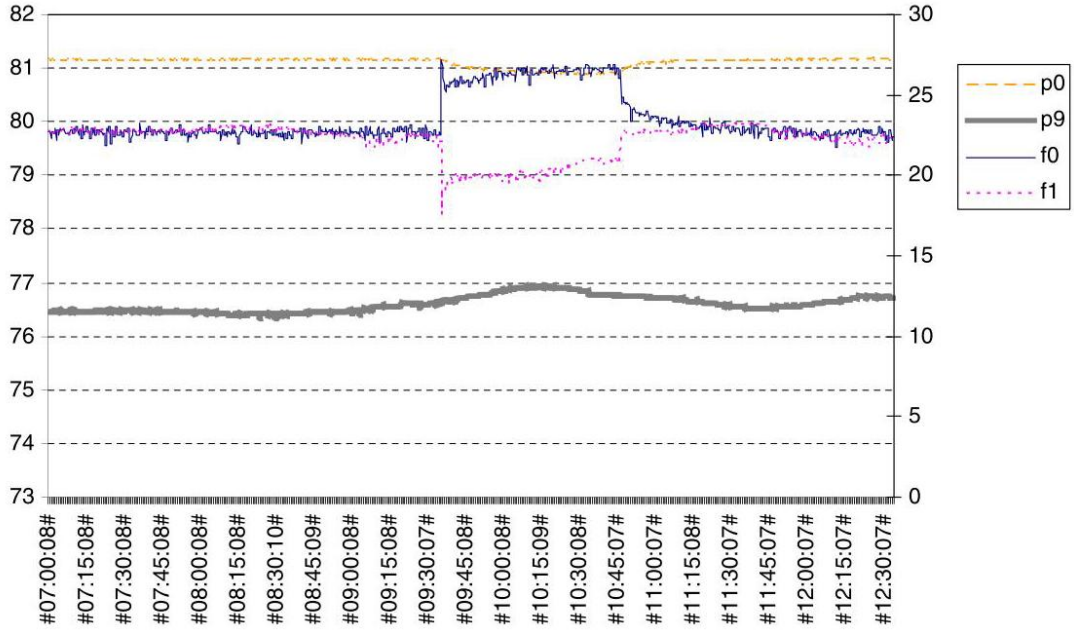


图 2. 入口和出口流量及压力读数。

表 1

计算得到的流量和压力差

	FlowDiff (f1 - f0)	Pressure Diff ($p_{average}(t) - p_{average}(t-1)$)	LeakSize t (h)
#07:00:18#	0.05	0	0
...
#10:04:08#	-6.15	-0.007	6.35
#10:04:18#	-6.15	0	6.38
#10:04:28#	-6.15	0	6.40
#10:04:39#	-6.1	0.0085	6.39
#10:04:49#	-6.1	0	6.35
...
#12:34:27#	-0.1	0	0

3.3. 前件和后件的参考点

每个前件使用的参考点数量决定了规则库的大小。如果数量太大，规则库中的规则会太多，后续的训练和推理过程会更复杂。如果数量太小，这些点可能无法覆盖前件属性的值域。对于传统规则库来说尤其如此。通常使用 5 - 9 个参考点。后件属性的参考点数量也与前件属性的参考点数量相当。在本文中，我们为 FlowDiff 使用 8 个点，分别是负大 (NL)、负中 (NM)、负小 (NS)、负非常小 (NVS)、零 (Z)、正小 (PS)、正中 (PM) 和正大 (PL)。即

$$A_1^k \in \{NL, NM, NS, NVS, Z, PS, PM, PL\} \quad (8)$$

类似地，我们为 PressureDiff 使用 7 个点，分别是 NL、NM、NS、Z、PS、PM 和 PL，即

$$A_2^k \in \{NL, NM, NS, Z, PS, PM, PL\} \quad (9)$$

对于后件属性 LeakSize，使用 5 个参考点：零 (Z)、非常小 (VS)、中等 (M)、高 (H) 和非常高 (VH)，即

$$\mathbf{D} = (D_1, D_2, D_3, D_4, D_5) = (\mathbf{Z}, \mathbf{VS}, \mathbf{M}, \mathbf{H}, \mathbf{VH}) \quad (10)$$

为了使用表 1 中所示的数据，上述为前件和后件属性定义的参考点是用语言术语表示的，需要进行量化。通过检查计算得到的 FlowDiff 和 PressureDiff 以及记录的 LeakSize 值，假设语言术语和数值之间存在以下等效关系，以便这些值大致覆盖相应属性的值域。

对于流量差 (FlowDiff)，假设

$$\begin{aligned} & \{ \text{NL} = -10 \\ & \text{NM} = -5 \\ & \text{NS} = -3 \\ & \text{NVS} = -1 \\ & \text{Z} = 0 \\ & \text{PS} = 1 \\ & \text{PM} = 2 \\ & \text{PL} = 3 \\ & \} \end{aligned} \quad (11)$$

对于压力差 (PressureDiff)，假设

$$\begin{aligned} & \{ \text{NL} = -0.01 \\ & \text{NM} = -0.005 \\ & \text{NS} = -0.002 \\ & \text{Z} = 0 \\ & \text{PS} = 0.002 \\ & \text{PM} = 0.005 \\ & \text{PL} = 0.01 \\ & \} \end{aligned} \quad (12)$$

对于泄漏尺寸 (LeakSize), 假设

$$\left\{ \begin{array}{l} Z = 0 \\ VS = 2 \\ M = 4 \\ H = 6 \\ VH = 8 \end{array} \right. \quad (13)$$

尽管参考值涵盖了历史数据中的属性值范围, 并且能够检测到尺寸高达管道吞吐量 30% 的泄漏, 但未来的泄漏尺寸有可能超过这一数值。当这种情况发生时, 泄漏尺寸将以 100% 的置信度被估计为 VH。

3.4. 规则

使用语言术语或其等效的参考数值, 用于泄漏检测和泄漏尺寸估计的传统规则之一可能如下所示:

如果流量差 (FlowDiff) 为 NL 且压力差 (PressureDiff) 为 NL

那么泄漏尺寸 (LeakSize) 为 VH。

然而, 这种传统的 IF - THEN 规则无法捕捉到前提条件和结论之间的连续关系。因此, 基于这些规则的专家系统可能无法实现较高的检测性能。

运用置信规则的概念, 传统规则可以扩展如下:

R_k : 如果流量差 (FlowDiff) 为 A_1^k 且压力差 (PressureDiff) 为 A_2^k

那么泄漏尺寸 (LeakSize) 为 $\{(Z, \beta_{1k}), (VS, \beta_{2k}), (M, \beta_{3k}), 56\}$ (14)

这里的 A_1^k, A_2^k 分别是式 (8) 和式 (9) 中定义的参考值。由于流量差 (FlowDiff) 被分为 8 个术语, 压力差 (PressureDiff) 被分为 7 个术语, 因此这两个前提条件有 56 种组合, 从而在规则库中总共产生 56 条规则。

初始置信规则可以通过以下四种方式建立:

(1) 从专家知识中提取规则, 形式可以是传统规则或置信规则。

表 2

增加泄漏对系统变量影响的总结

FlowDiff	Pressure Diff	LeakSize
Negative and constant	More negative	↑
Zero	More negative	↑
More negative	Negative and constant	↑
More negative	More negative	↑

(2) 通过检查历史数据提取规则。

(3) 如果有可用的, 使用先前用于类似管道泄漏检测的规则库。规则库可以是传统的规则库。

(4) 无任何先验知识的随机规则。

在我们的案例中, 没有先前的规则库作为起点。规则是通过检查数据并利用专家经验提取的, 并作为训练的起点。附录中的表 A1 列出了一位在管道泄漏检测方面有经验的专家提供的初始 56 条置信规则。例如, 表中的规则 8 表述如下:

R_8 : 如果流量差 (FlowDiff) 为负中 (NM) 且压力差 (PressureDiff) 为负小 (NL)

那么泄漏大小 (LeakSize) 为 $\{(H, 0.2), (VH, 0.8)\}$

在请专家提供规则时, 式 (11) 和 (12) 中所示语言术语的等效参考值尚未确定。因此, 这些规则可能在定性上是正确的, 即泄漏大小随流量差和压力差以正确的趋势变化 (如表 2 所示)。然而, 泄漏大小分布的置信度 (表 A1) 可能不准确。因此, 有必要进行进一步的训练来微调置信度, 以便在某种意义上提高或优化专家系统的性能。

3.5. 规则库的训练和测试

3.5.1. 训练数据

在泄漏试验期间, 以每个样本 10 s 的速率收集了 2008 个样本, 如图 2 所示。为了训练规则库, 选择了 500 个数据样本, 其中约一半是在泄漏期间收集的。它们是在上午 7 点至 7:33、上午 9:46 至 10:20 和上午 10:50 至 11:08 这三个时间段收集的数据 (图 2)。

3.5.2. 训练

使用第 2.3 节中定义的符号, 式 (5) 给出了需要训练的参数。更具体地说, 它们是

$$P = (\beta_{ik}, \theta_k, \delta_j; i = 1, \dots, N; k = 1, \dots, L; j = 1, \dots, T) \quad (15)$$

其中 $N = 5$, $L = 56$ 和 $T = 2$ 。

实际系统和基于规则的系统的输入都是由式 (6) 和 (7) 定义的流量差 FlowDiff(t) 和压力差 PressureDiff(t), 即

$$U(t) = (\text{FlowDiff}(t), \text{PressureDiff}(t))$$

观测输出 $\hat{O}(t)$ 是记录在数据文件中的受控 (或观测到的) 泄漏大小。模拟输出 $O(t)$ 使用下面步骤 5 中描述的推理方法计算。学习过程概述为以下七个步骤。该过程使用 MATLAB 实现。

步骤 1: 设置初始参数

初始置信度由专家给出, 并在表 A1 中列出。 θ_k 和 δ_j ($k = 1, \dots, 56; j = 1, 2$) 均假定为 1。

步骤 2: 转换输入

对于 500 个训练样本中的每一个, 输入值 [流量差 (FlowDiff(t)), 压力差 (PressureDiff(t))] 需要进行转换, 并根据式 (11) 和 (12) 中定义的参考值, 利用输入值与参考值的匹配置信度来表示。这种转换基于用于定量数据转换的基于规则的转换技术 (杨, 2001 年)。例如, 如果流量差 $(t) = -6.25$, 那么使用该属性的参考值, 它等效地转换为 $\text{FlowDiff}(t) = \{(NL, 0.25), (NM, 0.75)\}$, 因为 $-6.25 = NL \times 0.25 + NM \times 0.75 = -10 \times 0.25 + (-5) \times 0.75$ 。置信度 0.25 是输入流量差 $= -6.25$ 与参考值 $NL = -10$ 的匹配度。由于参考值 NL 在规则 1 至规则 7 中使用 (参见附录中的表 A1), 因此对于 $k = 1, 2, \dots, 7$, 式 (2) 中的 α_{1k} 为 0.25。类似地, 对于 $k = 8, 9, \dots, 14$, $\alpha_{1k} = 0.75$ 。

步骤 3: 计算规则激活权重

输入值与其属性参考值的匹配度 α_{ik} 用于公式 (2) 来计算第 k 条规则的激活权重 w_k ($k = 1, \dots, 56$)。

步骤 4: 合并激活的规则

采用证据推理 (ER) 方法 (Yang & Xu, 2002a) 来合并激活的规则。该 ER 方法在 IDS 软件中实现 (Yang & Xu, 2005)。使用 IDS 软件, 可以合并激活的规则以得到以下结果:

$$O(U(t)) = \{(D_j, \beta_j) \mid j = 1, \dots, 5\} \quad (16)$$

其中 β_i 由公式 (4) 给出, 且 $[D_1, D_2, D_3, D_4, D_5] = [Z, [VS, M, H, VH] = [0, 2, 4, 6, 8]$, 如公式 (13) 所定义。

步骤 5: 估计泄漏尺寸

在得到公式 (16) 所示的结果, 即模拟结果后, 泄漏尺寸 (LeakSize) 计算如下:

$$\text{EstimatedLeakSize}(t) = D_1\beta_1 + D_2\beta_2 + D_3\beta_3 + D_4\beta_4 + D_5\beta_5 \quad (17)$$

基于初始置信规则库, 使用公式 (17) 计算得到的估计泄漏尺寸如图 3 所示, 同时展示了训练数据中的观测泄漏尺寸。显然, 估计值与观测值匹配得不好。这意味着专家提供的初始规则库不够理想。

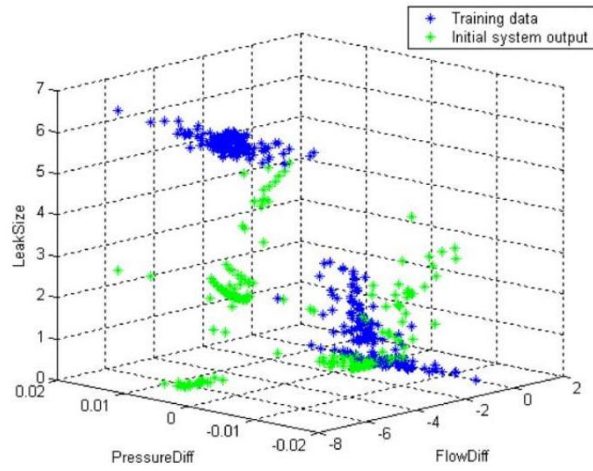


图 3. 训练数据和基于初始置信规则系统的输出。

步骤 6: 计算观测泄漏尺寸与估计泄漏尺寸之间的差值

观测泄漏尺寸与估计泄漏尺寸之间的差值 $\xi(P)$ 计算如下:

$$\xi(P) = \frac{1}{500} \times \sum_{t=1}^{500} (\text{ObservedLeakSize}(t) - \text{EstimatedLeakSize}(t))^2 \quad (18)$$

步骤 7: 找到一组新的参数 P , 使公式 (18) 定义的差值最小化

学习过程的目标是找到一组参数 P , 使观测泄漏尺寸与估计泄漏尺寸之间的差值最小化, 即:

$$\min_P \{\xi(P)\} \quad (19)$$

$$\text{s.t. } 0 \leq \beta_{ik} \leq 1 \quad (19a)$$

$$\sum_{i=1}^5 \beta_{ik} = 1 \quad (19b)$$

$$0 \leq \theta_k \leq 1 \quad (19c)$$

$$0 \leq \delta_j \leq 1 \quad (19d)$$

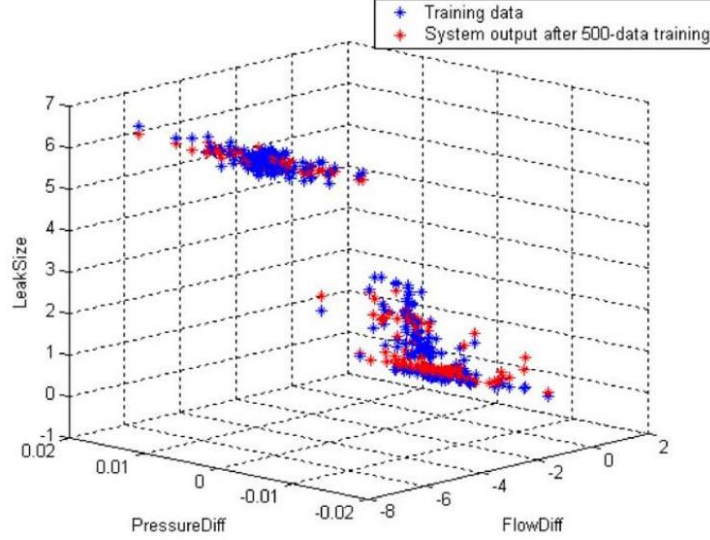


图 4. 训练数据和基于训练后置信规则系统的输出。

其中 $P = (\beta_{ik}, \theta_k, \delta_j; i = 1, \dots, 5; k = 1, \dots, 56; j = 1, 2)$ 。

上述问题使用 MATLAB 优化工具箱 (Coleman, Branch, & Grace, 1999) 解决。训练后的置信规则列于附录的表 A2 中。公式 (18) 所定义的观测泄漏尺寸与估计泄漏尺寸之间的差值小于 0.001。

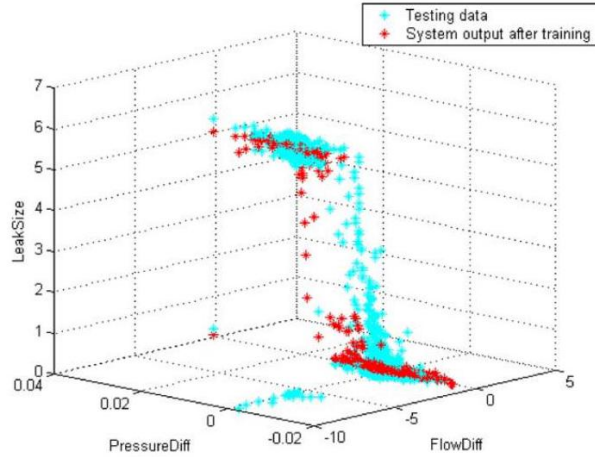


图 5. 测试数据以及训练后的基于置信规则的系统的输出。

图 4 显示，训练后的置信规则库能够紧密复现训练数据中流量差 (FlowDiff)、压力差 (PressureDiff) 和泄漏尺寸 (LeakSize) 之间的关系。

3.5.3. 测试

为了测试训练后的置信规则，使用了图 2 中所示的全部 2008 个样本。图 5 展示了相同前件值 (流量差和压力差) 下的观测泄漏尺寸和估计泄漏尺寸。这表明估计结果与观测结果非常接近。

图 6 在时间尺度上展示了观测泄漏尺寸和估计泄漏尺寸。它显示规则库能够清晰地检测到上午 9:36 左右发生、上午 10:47 左右结束的泄漏。

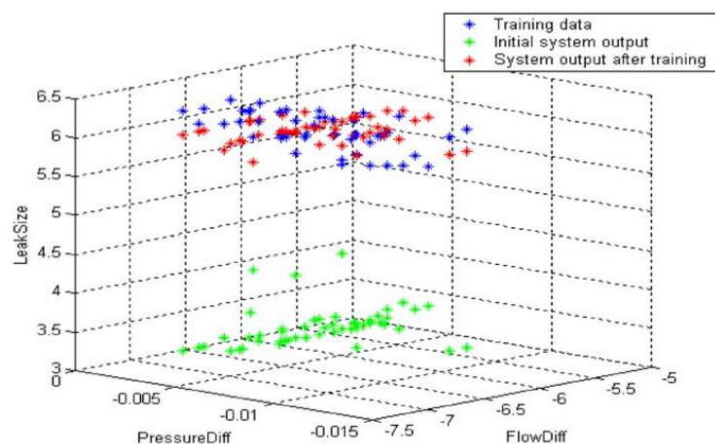


图 7. 随机初始规则库和训练后规则库生成的结果。

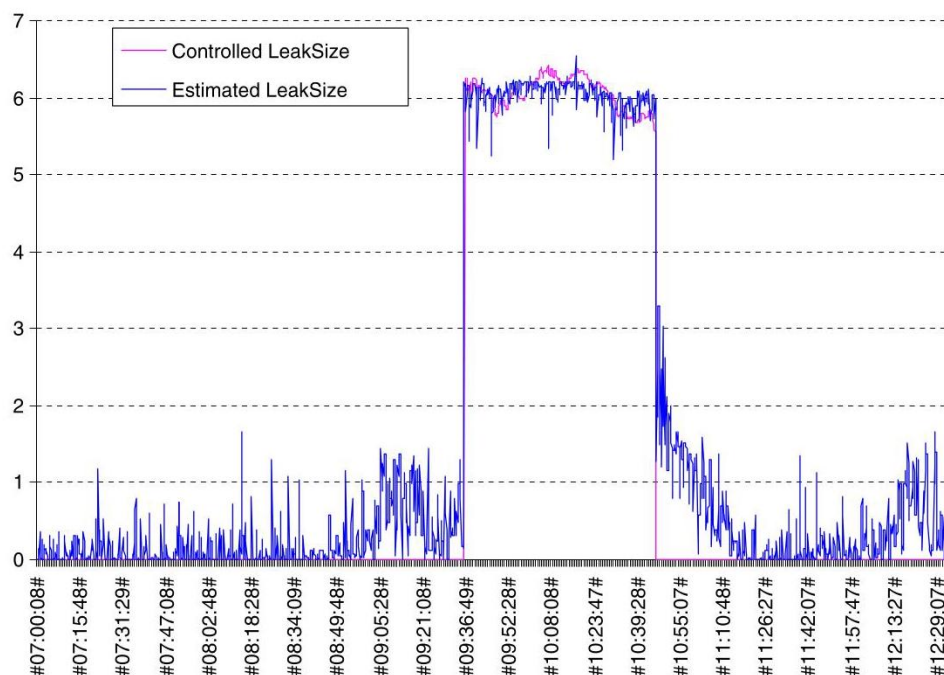


图 6. 测试数据以及训练后的基于置信规则的系统的输出。

还注意到，在无泄漏期间，估计的泄漏尺寸存在一定噪声，并非始终为零 (图 6)。这种噪声可能由管道中的湍流和动态变化，以及可能的仪器和数据通信误差引起。几乎所有管道运行数据都存在这种固有噪声，这给开发具有高可靠性和灵敏度的管道泄漏检测系统带来了重大挑战 (Bloom, 2004; Theakston, 2004; Theakston & Larnaes, 2002; Carpenter 等人, 2005)。防止噪声触发虚假泄漏警报的最常见方法是使用确认期。也就是说，在确认泄漏警报之前，估计的泄漏尺寸需要在指定时间段内持续为正值。

表 A1

专家提供的初始置信规则

Rule number	FlowDiff AND PressureDiff	LeakSize distribution $\{D_1, D_2, D_3, D_4, D_5\} = \{0, 2, 4, 6, 8\}$	LeakSize value
1	NL AND NL	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 1)\}$	8
2	NL AND NM	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 0.3), (D_5, 0.7)\}$	7.4
3	NL AND NS	$\{(D_1, 0), (D_2, 0), (D_3, 0.2), (D_4, 0.8), (D_5, 0)\}$	5.6
4	NL AND Z	$\{(D_1, 0), (D_2, 0), (D_3, 0.8), (D_4, 0.2), (D_5, 0)\}$	4.4
5	NL AND PS	$\{(D_1, 0.65), (D_2, 0.35), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.7
6	NL AND PM	$\{(D_1, 0.85), (D_2, 0.15), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.3
7	NL AND PL	$\{(D_1, 0.95), (D_2, 0.05), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.1
8	NM AND NL	$\{(D_1, 0), (D_2, 0), (D_3, 0.1), (D_4, 0.9), (D_5, 0)\}$	5.8
9	NM AND NM	$\{(D_1, 0), (D_2, 0), (D_3, 0.7), (D_4, 0.3), (D_5, 0)\}$	4.6
10	NM AND NS	$\{(D_1, 0), (D_2, 0.7), (D_3, 0.3), (D_4, 0), (D_5, 0)\}$	2.6
11	NM AND Z	$\{(D_1, 0), (D_2, 0.9), (D_3, 0.1), (D_4, 0), (D_5, 0)\}$	2.2
12	NM AND PS	$\{(D_1, 0.8), (D_2, 0.2), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.4
13	NM AND PM	$\{(D_1, 0.9), (D_2, 0.1), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.2
14	NM AND PL	$\{(D_1, 0.99), (D_2, 0.01), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.02
15	NS AND NL	$\{(D_1, 0), (D_2, 0), (D_3, 0.4), (D_4, 0.6), (D_5, 0)\}$	5.2
16	NS AND NM	$\{(D_1, 0), (D_2, 0), (D_3, 0.8), (D_4, 0.2), (D_5, 0)\}$	4.4
17	NS AND NS	$\{(D_1, 0), (D_2, 0.3), (D_3, 0.6), (D_4, 0.1), (D_5, 0)\}$	3.6
18	NS AND Z	$\{(D_1, 0.1), (D_2, 0.7), (D_3, 0.2), (D_4, 0), (D_5, 0)\}$	2.2
19	NS AND PS	$\{(D_1, 0.7), (D_2, 0.3), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.6
20	NS AND PM	$\{(D_1, 0.9), (D_2, 0.1), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.2
21	NS AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
22	NVS AND NL	$\{(D_1, 0), (D_2, 0.1), (D_3, 0.4), (D_4, 0.5), (D_5, 0)\}$	4.8
23	NVS AND NM	$\{(D_1, 0), (D_2, 0.8), (D_3, 0.2), (D_4, 0), (D_5, 0)\}$	2.4
24	NVS AND NS	$\{(D_1, 0.2), (D_2, 0.7), (D_3, 0.1), (D_4, 0), (D_5, 0)\}$	1.8
25	NVS AND Z	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
26	NVS AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
27	NVS AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
28	NVS AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
29	Z AND NL	$\{(D_1, 0), (D_2, 0.4), (D_3, 0.6), (D_4, 0), (D_5, 0)\}$	3.2
30	Z AND NM	$\{(D_1, 0.2), (D_2, 0.7), (D_3, 0.1), (D_4, 0), (D_5, 0)\}$	1.8
31	Z AND NS	$\{(D_1, 0.4), (D_2, 0.6), (D_3, 0), (D_4, 0), (D_5, 0)\}$	1.2
32	Z AND Z	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
33	Z AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
34	Z AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
35	Z AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
36	PS AND NL	$\{(D_1, 0), (D_2, 0.8), (D_3, 0.2), (D_4, 0), (D_5, 0)\}$	2.4
37	PS AND NM	$\{(D_1, 0.8), (D_2, 0.2), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.4
38	PS AND NS	$\{(D_1, 0.95), (D_2, 0.05), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.1
39	PS AND Z	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
40	PS AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
41	PS AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
42	PS AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
43	PM AND NL	$\{(D_1, 0.1), (D_2, 0.9), (D_3, 0), (D_4, 0), (D_5, 0)\}$	1.8
44	PM AND NM	$\{(D_1, 0.3), (D_2, 0.7), (D_3, 0), (D_4, 0), (D_5, 0)\}$	1.4
45	PM AND NS	$\{(D_1, 0.85), (D_2, 0.15), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.3
46	PM AND Z	$\{(D_1, 0.98), (D_2, 0.02), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.04
47	PM AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
48	PM AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
49	PM AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
50	PL AND NL	$\{(D_1, 0.9), (D_2, 0.1), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.2
51	PL AND NM	$\{(D_1, 0.99), (D_2, 0.01), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.02
52	PL AND NS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
53	PL AND Z	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
54	PL AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
55	PL AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
56	PL AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0

确认期的长度会因所用仪器的质量、估计的泄漏尺寸、管道中输送的物料以及管道是处于稳定、变化、启动还是关闭状态而有所不同。也可以建立、训练另一个基于置信规则的系统来检测管道运行状态并确定确认期的长度，本文对此不作讨论。在我们的案例中，如果将确认期设置为 24 min，那么由数据噪声引起的所有虚假警报将被完全消除。

表 A2

经过 500 个数据训练后的训练置信规则 (假设初始规则权重相等)

Rule number	Trained rule weight	FlowDiff AND PressureDiff	LeakSize distribution $\{D_1, D_2, D_3, D_4, D_5\} = \{0, 2, 4, 6, 8\}$	LeakSize value
1	0.90	NL AND NL	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 1)\}$	8
2	0.90	NL AND NM	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 0.3), (D_5, 0.7)\}$	7.4
3	0.50	NL AND NS	$\{(D_1, 0), (D_2, 0), (D_3, 0.2), (D_4, 0.8), (D_5, 0)\}$	5.6
4	0.50	NL AND Z	$\{(D_1, 0), (D_2, 0), (D_3, 0.8), (D_4, 0.2), (D_5, 0)\}$	4.4
5	0.50	NL AND PS	$\{(D_1, 0.65), (D_2, 0.35), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.7
6	0.80	NL AND PM	$\{(D_1, 0.85), (D_2, 0.15), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.3
7	0.90	NL AND PL	$\{(D_1, 0.95), (D_2, 0.05), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.1
8	0.90	NM AND NL	$\{(D_1, 0), (D_2, 0), (D_3, 0.1), (D_4, 0.9), (D_5, 0)\}$	5.8
9	0.80	NM AND NM	$\{(D_1, 0), (D_2, 0), (D_3, 0.7), (D_4, 0.3), (D_5, 0)\}$	4.6
10	0.50	NM AND NS	$\{(D_1, 0), (D_2, 0.7), (D_3, 0.3), (D_4, 0), (D_5, 0)\}$	2.6
11	0.50	NM AND Z	$\{(D_1, 0), (D_2, 0.9), (D_3, 0.1), (D_4, 0), (D_5, 0)\}$	2.2
12	0.50	NM AND PS	$\{(D_1, 0.8), (D_2, 0.2), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.4
13	0.90	NM AND PM	$\{(D_1, 0.9), (D_2, 0.1), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.2
14	0.90	NM AND PL	$\{(D_1, 0.99), (D_2, 0.01), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.02
15	0.80	NS AND NL	$\{(D_1, 0), (D_2, 0), (D_3, 0.4), (D_4, 0.6), (D_5, 0)\}$	5.2
16	0.80	NS AND NM	$\{(D_1, 0), (D_2, 0), (D_3, 0.8), (D_4, 0.2), (D_5, 0)\}$	4.4
17	0.50	NS AND NS	$\{(D_1, 0), (D_2, 0.3), (D_3, 0.6), (D_4, 0.1), (D_5, 0)\}$	3.6
18	0.50	NS AND Z	$\{(D_1, 0.1), (D_2, 0.7), (D_3, 0.2), (D_4, 0), (D_5, 0)\}$	2.2
19	0.50	NS AND PS	$\{(D_1, 0.7), (D_2, 0.3), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.6
20	0.50	NS AND PM	$\{(D_1, 0.9), (D_2, 0.1), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.2
21	1.00	NS AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
22	0.77	NVS AND NL	$\{(D_1, 0.02), (D_2, 0.11), (D_3, 0.39), (D_4, 0.48), (D_5, 0)\}$	4.66
23	0.36	NVS AND NM	$\{(D_1, 0.10), (D_2, 0.78), (D_3, 0.12), (D_4, 0), (D_5, 0)\}$	2.04
24	0.30	NVS AND NS	$\{(D_1, 0.36), (D_2, 0.65), (D_3, 0.10), (D_4, 0), (D_5, 0)\}$	1.75
25	1.00	NVS AND Z	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
26	1.00	NVS AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
27	1.00	NVS AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
28	1.00	NVS AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
29	0.64	Z AND NL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
30	0.88	Z AND NM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
31	0.55	Z AND NS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
32	1.00	Z AND Z	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
33	1.00	Z AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
34	1.00	Z AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
35	1	Z AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
36	0.77	PS AND NL	$\{(D_1, 0.39), (D_2, 0.61), (D_3, 0), (D_4, 0), (D_5, 0)\}$	1.22
37	0.64	PS AND NM	$\{(D_1, 0.90), (D_2, 0.1), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.2
38	0.57	PS AND NS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
39	1.00	PS AND Z	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
40	1.00	PS AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
41	1.00	PS AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
42	1.00	PS AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
43	0.90	PM AND NL	$\{(D_1, 0.1), (D_2, 0.9), (D_3, 0), (D_4, 0), (D_5, 0)\}$	1.8
44	0.50	PM AND NM	$\{(D_1, 0.3), (D_2, 0.7), (D_3, 0), (D_4, 0), (D_5, 0)\}$	1.4
45	0.50	PM AND NS	$\{(D_1, 0.85), (D_2, 0.15), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.3
46	0.50	PM AND Z	$\{(D_1, 0.98), (D_2, 0.02), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.04
47	1.00	PM AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
48	1.00	PM AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
49	1.00	PM AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
50	0.90	PL AND NL	$\{(D_1, 0.9), (D_2, 0.1), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.2
51	0.70	PL AND NM	$\{(D_1, 0.99), (D_2, 0.01), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0.02
52	1.00	PL AND NS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
53	1.00	PL AND Z	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
54	1.00	PL AND PS	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
55	1.00	PL AND PM	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0
56	1.00	PL AND PL	$\{(D_1, 1), (D_2, 0), (D_3, 0), (D_4, 0), (D_5, 0)\}$	0

3.6. 随机初始规则训练

在上一节中，初始规则由专家给出。为了测试基于置信规则的系统的学习能力，使用了多个随机生成的规则库作为学习的起点。研究表明，初始随机规则可以被训练以表征训练数据中的关系，如图 7 所示，其中初始系统输出由随机置信规则库生成。训练后，系统输出与观测输出非常匹配。置信规则库的这种学习能力使其适用于各种情况，无论是否有专家知识可用。

4. 结论

本文描述了将基于置信规则的系统应用于管道泄漏检测的可行性研究。研究表明，通过使用最优学习方法 (Yang 等人, 2005b) 来训练置信规则，基于置信规则的系统可以从管道运行数据中学习泄漏尺寸与管道流量和压力读数之间的关系。研究还表明，学习可以从一个随机的规则库开始，因此无需提供先验知识。

对于基于质量平衡原理的管道泄漏检测系统而言，识别流量差与压力变化之间的关系非常重要，并且需要进行大量的参数调整。这是一个耗时的过程，并且具有管道特异性。随着管道和仪器的老化，最优参数也可能随时间发生变化。基于置信规则的系统的自学习能力可以显著减少参数调整时间，并提高系统的性能。

传统规则库是置信规则库的一种特殊情况，也可用于此目的。然而，由于它无法对泄漏尺寸与管道仪器读数之间的连续关系进行建模，因此很难准确估计泄漏尺寸。

神经网络 (Gallant, 1993) 也是一种常用于对不同变量之间因果关系进行建模的工具。在基于置信规则的系统中，虽然人类专家知识用于构建大致正确的置信规则库，但如果有系统输入 - 输出数据，最优学习机制可以帮助微调系统性能。因此，我们认为，与黑箱系统给出的建议相比，经过微调的逻辑规则推理更易被人类用户接受，因为这种推理是可理解的、能提供解释，并且可以通过人工检查进行验证。它还能增强对系统的信心，并有助于发现重要的关系和特征组合。

为了将基于置信规则的系统应用于实时在线泄漏检测，需要辅助置信规则库，以便在线做出智能判断，判断仪器是否故障、管道是正在启动、关闭还是处于稳定运行状态，以及仪器读数是否漂移。这些判断有助于动态调整确认周期的长度，并校正输入到置信规则中的实际流量和压力差。借助此类基于辅助置信规则的系統，我们期望新的泄漏检测系统的灵敏度、可靠性和响应时间能够得到改善。然而，在这一点得到证实之前，还需要进行进一步的研究和可行性分析。

致谢

这项工作是由英国工程与物理科学研究委员会资助的项目的一部分，资助编号为:GR/S85498/01 和 GR/S85504/01。作者也感谢香港 KC 王教育基金会的支持。

附录

参见表 A1 和表 A2。

参考文献

- Abhulimen, K. E., & Susu, A. A. (2004). Liquid pipeline leak detection system: model development and numerical simulation. *Chemical Engineering Journal*, 97(1), 47-67.
- Bloom, D. (2004). Non-intrusive system detects leaks using mass measurement. *Pipeline & Gas Journal*, 231(7), 20-21.

- Buchberger, S. G., & Nadimpalli, G. (2004). Leak estimation in water distribution systems by statistical analysis of flow readings. *Journal of Water Resources Planning and Management - ASCE*, 130(4), 321- 329.
- Carpenter, P. S., Nicholas, E., & Henrie, M. (2005). Method gives realistic analysis of leak-detection systems. *Oil & Gas Journal*, 103(11), 53-58.
- Coleman, T., Branch, M. A., & Grace, A. (1999). Optimization toolbox - for use with MATLAB. The Mathworks Inc.
- Gallant, S. I. (1993). Neural network learning and expert systems. MIT Press.
- Gao, Y., Brennan, M. J., Joseph, P. F., Muggleton, J. M., & Hunaidi, O. (2005). On the selection of acoustic/vibration sensors for leak detection in plastic water pipes. *Journal of Sound and Vibration*, 283(3-5), 927-941.
- Hodges, J., Bridges, S., Sparrow, C., Wooley, B., Tang, B., & Jun, C. (1999). The development of an expert system for the characterization of containers of contaminated waste. *Expert Systems with Applications*, 17, 167-181.
- Lee, M. R., Lee, J. H., & Park, Y. C. (2004). Development of PC-based leak detection system using acoustic emission technique. *Advances in Nondestructive Evaluation, Part 1-3, Key Engineering Materials(270- 273)*, 525-530.
- Parson, S. (1996). Current approaches to handling imperfect information in data and knowledge bases. *IEEE Transactions on Knowledge and Date Engineering*, 8(3), 353-372.
- Rougier, J. Probabilistic Leak Detection in Pipelines Using the Mass Imbalance Approach. *Journal of Hydraulic Research*. Available from <http://maths.dur.ac.uk/stats/people/jcr/newSSdesign.pdf>, visited on 23 June [forthcoming].
- Shafer, G. (1976). A mathematical theory of evidence. Princeton, NJ: Princeton University Press.
- Sun, R. (1995). Robust reasoning: integrating rule-based and similarity-based reasoning. *Artificial Intelligence*, 75, 241-295.
- Theakston, J. (2004). Good specifications can result in useful software-based leak detection. *Pipeline & Gas Journal*, 231(3), 43-48.
- Theakston, J., & Larnaes, G. (2002). Selecting and installing a software-based leak detection system. *Pipeline & Gas Journal*, 229(10), 52-55.
- Wang, Y. M., Yang, J. B. & Xu, D. L. (2005). Environmental impact assessment using the evidential reasoning approach. *European Journal of Operational Research* [in press].
- Yang, J. B. (2001). Rule and utility based evidential reasoning approach for multi-attribute decision analysis under uncertainties. *European Journal of Operational Research*, 131, 31-61.
- Yang, J. B., & Sen, P. (1994). A general multi-level evaluation process for hybrid MADM with uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(10), 1458-1473.
- Yang, J. B., & Singh, M. G. (1994). An evidential reasoning approach for multiple attribute decision making with uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(1), 1-18.
- Yang, J. B., & Xu, D. L. (2002a). On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 32(3), 289-304.
- Yang, J. B., & Xu, D. L. (2002b). Nonlinear information aggregation via evidential reasoning in multi-attribute decision analysis under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 32(3), 376-393.
- Yang, J. B. & Xu, D. L. (2005) IDS Multicriteria Assessor, Version 2.1. Available from www.e-ids.co.uk.
- Yang, J. B., Liu, J., Wang, J., Sii, H. S., & Wang, H. W. (2005a). A generic rule-base inference methodology using the evidential reasoning approach - RIMER. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* [in press].

Yang, J. B., Liu, J., Xu, D. L., Wang, J., & Wang, H. W. (2005b). Optimal learning method for training belief rule based systems. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* [under review].