

农业远程诊断 方法和应用

◎ 周国民 丘 耘 周义桃 著



中国农业科学技术出版社

国家“863”项目资助

农业远程 诊断方法和应用

周国民 丘 耘 周义桃 著

中国农业科学技术出版社

图书在版编目 (CIP) 数据

农业远程诊断方法和应用/周国民等著. —北京:
中国农业科学技术出版社, 2006
ISBN 7-80233-062-9

I. 农... II. 周... III. 计算机网络—应用—
农业技术—研究 IV. S-39

中国版本图书馆 CIP 数据核字 (2006) 第 091991 号

责任编辑: 梅 红

责任校对: 贾晓红

出版发行: 中国农业科学技术出版社

北京市中关村南大街 12 号 邮编: 100081

电话: (010) 62189012 (编辑室) 传真: (010) 68975144

(010) 68919704 (发行部)

经 销 者: 新华书店北京发行所

印 刷 者: 北京雅艺彩印有限公司

开 本: 850mm × 1168mm 1/32

印 张: 8.125

字 数: 250 千字

版 次: 2006 年 11 月第 1 版

印 次: 2006 年 11 月第 1 次印刷

印 数: 1 ~ 1500 册

定 价: 20.00 元

版权所有 · 翻印必究

序

20 世纪 90 年代后期，信息通讯技术的快速发展，为农业远程诊断技术的研究应用提供了有力支撑。农业远程诊断平台可以在农业科研机构的专家与广大基层农户之间架起一个信息沟通与服务的桥梁，有效解决农村基层对农业科学知识与技术的需求问题，促进农业产业化和农村信息化的发展。当前，我国正在大力建设社会主义新农村，农业远程诊断技术的广泛应用对于实现“信息惠农”、“信息惠民”理念，培养和造就一批知识型农民，促进新农村建设具有重要意义。

《农业远程诊断方法和应用》是作者所在的研究团队开展国家 863 计划课题“网络化实时农业远程诊断模型与交互式关键平台技术研究”的成果系统总结，是我国农业科技领域比较系统论述农业问题远程诊断的著作。该书内容丰富，特色鲜明，联系实际，注重应用，以“方法基础——技术研究——实践应用”为主线，系统阐述了每种诊断方法的技术原理、应用现状、应用范围以及如何与相关技术集成等问题。同时还具体阐述了如何在网络环境中应用专家系统方法、事例库方法、神经网络等方法开发一个实际的农业远程诊断平台，特别是对于基于二叉树判定型知识的可视化编辑与应用、基于 SDD 算法的事例库推理机的设计、基于 XML 的网络化专家系统的设计与实现等进行了系统详实的阐述，具有很好的借鉴参考价值。

本书作者是活跃在我国农业信息技术研究领域的青年专家，

具有较好的信息技术与农学复合型的知识结构，思维敏捷，工作扎实，态度严谨，从本书论述中也能领略其科研实践中的创新能力。

国家农业信息化工程技术研究中心	主 任
农业部农业信息技术重点开放实验室	主 任
北京农业信息化学会	理事长
赵春江	研究员
2006 年 8 月 8 日	于北京

前 言

我国幅员辽阔，大多数农民生活在交通不便的农村，而掌握先进技术的农业科研机构、农业专家大多集中在大中城市。因此，仅靠传统的派“支农小分队”等形式，显然已不能适应广大基层农民对农业科学与技术信息服务的需求。近年来，我们在考虑国内外有关农业技术推广发展趋势时，注意到国外许多国家已经出现了利用新兴的信息技术和网络技术构建远程诊断平台的趋势。利用远程诊断平台，专家和基层农业生产人员之间有了一个好的信息沟通平台，由此萌发了在我国农业领域开展农业远程诊断研究的想法。从2002年开始，中国农业科学院科技文献信息中心（现更名为中国农业科学院农业信息研究所）多媒体技术研究室在国家“863”计划的支持下，开展以大田作物病虫害诊断为主要内容的远程咨询平台的研究。本书是国家“863”项目“网络化实时农业远程诊断模型与交互式关键平台技术研究”课题的研究成果之一。

本书以“方法基础——技术研究——实践应用”为主线，系统阐述了不同诊断方法的技术原理、应用现状、应用范围以及如何与相关技术集成等问题。同时还具体阐述了如何在网络环境中应用专家系统方法、事例库方法、神经网络等方法开发一个实际的农业远程诊断平台。最后还列举了农业远程诊断平台在黑龙江、重庆、江苏、山东、河北、辽宁6个省（市）进行示范应用的情况，为今后建立类似的农业远程诊断平台提供了经验。

由于国内农业领域讲述远程诊断的专著不多，因此本书可以借

鉴的材料非常少，再加上农业远程诊断的研究刚刚起步，还存在一些问题有待系统地研究，相关技术还不是很成熟，因此我们面临的各种各样的困难非常多。但本书从研究到编撰过程中一直得到中国农业科学院农业信息研究所的领导和有关部门的大力支持，并得到国家“863”项目“网络化实时农业远程诊断模型与交互式关键平台技术研究”的课题资助。为此，我们对他们远见卓识的行动表示衷心的感谢。

本书是项目组全体成员共同辛勤工作的结晶，集中了撰写人员、专家顾问以及农业生产基层单位工作人员的智慧，是多次反复讨论、交流、修改的结果。其中，农业远程诊断概述、大田作物病虫害远程诊断系统由周国民执笔，网络技术由周义桃执笔，专家系统方法由丘耘执笔，网络技术由樊景超、李蕾执笔，事例库方法、ExpertMap 系统由李蕾、邵世磊、樊景超执笔，神经网络方法、基于神经网络的水稻病虫害识别系统由王剑执笔。全书最后由周国民、周义桃、丘耘修改和统稿。

本书在编写过程中，得到了赵春江研究员的大力支持与帮助，并为本书做序，在此表示衷心的感谢。

感谢中国农业科学技术出版社编辑在本书出版工作中给予的大力支持与帮助。

我们真诚地欢迎大家批评指正。

编 者

2006 年 11 月

目 录

农业远程诊断概述

- 一、诊断理论模型 (1)
- 二、诊断技术的方法与趋势 (3)
- 三、农业远程诊断 (10)

专家系统方法

- 一、专家系统的现状与发展 (12)
- 二、专家系统原理 (19)
- 三、CLIPS 系统 (35)
- 四、二叉树型的知识编辑方法 (54)

事例库方法

- 一、事例推理 (63)
- 二、事例推理系统及案例 (74)
- 三、事例的语义检索算法研究 (90)

神经网络方法

- 一、神经网络的应用现状及发展 (107)
- 二、神经网络原理 (117)
- 三、BP 算法 (120)

网 络 技 术

- 一、视频数据网络传输技术 (129)
- 二、短信网关开发技术 (134)

三、Web 程序设计技术	(139)
--------------------	-------

网 格 技 术

一、网格及网格技术概述	(146)
二、网格技术	(157)
三、网格应用	(169)

大田作物病虫害远程诊断系统

一、系统原理	(180)
二、系统的结构与功能	(185)
三、系统的特点与应用	(199)

ExpertMap 系统

一、系统原理	(204)
二、系统结构	(210)
三、系统的主要功能及在线咨询流程	(213)
四、系统的主要特点	(220)

基于神经网络的水稻病虫害识别系统

一、系统简介	(224)
二、系统总体框架和主要功能	(225)
三、主要特点	(239)

参考文献	(242)
------------	-------

农业远程诊断概述

农业远程诊断是一个集农业信号远程采集、处理、分析及评判的过程,通过这个过程将农业系统中某个子系统的运行状态分为正常和异常两类,并进一步找出异常的信号样本所属的类别,因此农业远程诊断问题又可以看成是模式分类与识别问题。

一、诊断理论模型

诊断技术的研究已历经几十年,但是研究成果尚不足以形成完整的系统理论体系;另一方面,应用技术的研究也并不完善,目前的工作一般都是集中在给定系统中特定诊断的技术实现,而诸如多变量、多输入与多输出系统、大系统、严重非线性系统等的诊断技术,以及对噪声和模型的不确定性对象具有鲁棒智能诊断的系统研究等,有些正在开展,有些则尚有难度。

在某种病态下,农业系统的性能明显低于正常水平,难以完成系统规定的或预定的任务。诊断的目的与任务首先就是要检测并确认这种状态或提前预报这种状态的出现。大部分的被诊断对象都可归纳为如下的模型:

$$\Sigma: Y(t) = f[X(t), U(t)]$$

式中: $Y(t)$ ——系统输出 m 为向量, $Y(t) \in L_m$ (输出空间);

f ——一个映射;

$X(t)$ —— n 为状态向量, $X(t) \in L_n$ (系统的状态空间);

$U(t)$ ——系统激励输入是 l 维向量, $U(t) \in L_l$ (输入空间)。

诊断系统状态向量 $X(t)$ 的选择与控制系统状态向量的选择不一定相同。在控制系统设计的时候,往往选择那些容易观测或检测的状态及反映系统运行特征的变量作为系统的状态向量。而在诊断系统设计时,一般总是选择那些反映系统内部特征,尤其是能表述系统内部子系统和部件工作状态的变量作为状态向量,希望这些状态向量能充分表达和描述系统病态时的特征,以方便诊断功能的实现。

系统状态向量 $X(t)$ 的特征只能在系统的采样信息 $I(t) = [Y(t), U(t)]$ 中获取。设 R 是 $X(t)$ 所有的特征集合, Q 是一个可数集合, S 是一个可数集合(所有的推理控制策略), δ_1 、 δ_2 是状态转移函数, b 是一个空白符号集合。所谓诊断问题 P 就是根据系统信息 $I(t)$ 确定 S 的控制状态,由状态转移函数 δ_1 实现由 $I(t)$ 到 R , 然后由 δ_2 实现 R 至 Q 的一个映射。映射的结果,可能是一个空白符。由于诊断对象的时序关系和病态之间的耦合关系,推理的初始出发状态为 q_0 , 终了状态为 q_f 。诊断过程可归结为如下的九元有限诊断自动机问题:

$$P = \{I, R, Q, \delta, S, D, b, q_0, q_f\}$$

式中: $\delta = \{\delta_1, \delta_2\}$

若依据一定的系统信息,并非所有的病态都是可诊断的。一个系统在给定的输出条件之下,系统内部某些部件的病态并不一定影响系统的输出,而系统诊断功能的完成是在系统输入输出信息的基础上,根据信息变化的异常特征来实现的。如系统中某些用来改善系统动态特性的子系统病态时,在一定的条件之下并不(或几乎不)影响系统的输出或功能。因此存在一个可诊断与不可诊断的问题。系统的状态能够在系统中反映出来,则系统是输出能观测的,也就是系统的输出中含有系统状态的特征,即系统相应的状态能诊断;但是有些系统状态能诊断却并非能观测。因此系统状态的能诊断性反映了系统的可利用信息集合 $I(t) = [Y(t), U(t)]$

与相应的状态特征集合 R 之间的相关程度, 即 $I(t) = [Y(t), U(t)]$ 对 R 的表征能力。

二、诊断技术的方法与趋势

(一) 诊断技术的方法

与人工智能技术发展紧密相连的智能诊断技术和方法主要有: 基于规则的方法、基于模型知识的方法、基于神经网络的方法、基于案例的方法、基于行为的方法等。这些方法虽各有特点和优势, 但均存在各自的局限性。

1. 基于规则的诊断

基于规则的诊断方法是根据被诊断系统的专家以往诊断的经验, 将其归纳成规则, 并运用经验规则通过推理来进行病态诊断。该方法具有诊断过程简便、快速等优点, 由于该方法属于反演推理, 因而不是一种确保唯一性的推理形式, 存在着知识获取瓶颈。复杂系统所观测到的症状与所对应诊断之间的联系相当复杂, 专家经验归纳成规则往往不是唯一的且有相当难度。因此该方法不适用于复杂系统或新的及尚无经验的系统的病态诊断。基于规则的方法对于诊断结论除了重复被采用的规则外, 无法做出进一步解释, 通常只能诊断单个病态, 难以诊断多重病态。

2. 基于模型知识的诊断

基于模型知识的诊断方法是运用被诊断系统的运行模型和病态模型, 由模型获得的预测形态和所测量观察的形态之间的差异, 计算出被诊断系统的最小诊断。其原理是先通过测量获取所观察的形态, 后基于系统输入和所有组件的正常状态模型获得预测形态, 根据这两种形态的差异 ATMS (Assumption based Truth Maintenance System) 计算出最小冲突集, 再由最小冲突集计算出最小诊断。最

小诊断就是关于病态元件的假设。基于模型的诊断方法具有不依赖被诊断系统的诊断训练例子和诊断经验,因而适用于新的及尚无经验的系统的病态诊断,能够诊断多重病态并能对诊断结论进行解释。对于大型被诊断系统,基于模型的诊断方法计算量大;诊断结果给出许多关于病态元件的假设,诊断过程需进行计算、再多次迭代才能找出病态。近年来发展了基于经验知识和模型知识相结合的诊断推理方法,如 Gallanti 和 Fink 提出的集成诊断模型, Peng 的层次因果模型等。

3. 基于神经网络的诊断

基于神经网络的智能病态诊断研究集中在三方面:一是从模式识别角度应用神经网络作为分类器进行病态诊断;二是从预测角度应用神经网络作为动态预报模型进行病态预测;三是从知识处理角度建立基于神经网络的诊断专家系统。

(1) 病态诊断的神经网络

基于神经网络模型的智能诊断,将病态诊断问题视为模式分类和识别问题,利用神经网络优越的模式分类性能进行诊断。用于病态模式识别的 ANN 模型,按学习方式可分为有监督学习模型和无监督学习模型两大类,前者主要包括 BP 网和径向基函数(RBF)网;后者主要包括自适应共振(ART)网和自组织特征映射(SOM)网。

从已有研究结果看,监督学习模型有很好的推广能力,用于病态模式识别效果较好;训练好的 BP 网和 RBF 网计算速度快,消耗内存少,可用于实时监测与诊断。但是这类模型要求学习样本具有一定的致密性、遍历性和相容性,工程实际中要获得这样的样本有时还很困难。此外,知识的分布式表达也使其失去了对诊断结果必要的解释能力。

ART 网和 SOM 网均属于无监督竞争学习的自组织模型,ART 网可以在线学习,边学习边记忆,给网络一批输入样本,它自动形

成一组分类模式。当一个新的输入不能归入已形成的模式类时，网络又自动形成一个新的模式类；若新的输入在已形成的模式类中可以找到一个相似的类，则这个输入划入该模式类，同时网络向更接近这个输入的方向作调整。SOM 网采用离线方式学习，能很好地进行特征提取，适用于作最邻近分类器。目前，自组织模型应用于病态诊断的尚少，它是一类次优的模式分类器，推广性能不如监督学习模型的好，但它具有的自组织和自适应特性正引起人们广泛关注。

(2) 病态预测的神经网络

病态预测的神经网络主要以两种方式实现预测功能：一是以神经网络（如 BP 网络）作为函数逼近器，对预测对象的某参数进行拟和预测；二是根据输入或输出间的动态关系，建立动态模型而进行病态预测。

动态神经网络的预测是一个对动态时序建模的过程。动态神经网络是由于神经元连接中存在时延反馈形成的，又称为回归网络，具有记忆和联想功能。它又可分为全回归网络和具有局部信息反馈结构的局部回归网络。这些网络的一个共同特点是其输出不仅取决于当前输入，网络本身具有相应的动态结构，其预测是动态预测，因而在实际的非线性动态系统的建模和预测中得到了成功的应用。但是由于动态神经网络在结构上远比前馈网络结构复杂，其样本训练也较困难，因此合理地降低网络结构的复杂性，简化网络的学习算法将是实际应用中待解决的问题。

(3) 基于神经网络的诊断专家系统

神经网络与专家系统的结合主要有两种策略：一是使用神经网络来构造专家系统，即把传统专家系统基于符号的推理变成基于数值运算的推理，以提高专家系统的执行效率，并解决专家系统的自学习问题；二是将神经网络理解为一类知识源的表达与处理模式，这些模式与其他知识表达方式如规则、框架等一起来表达领域专家

的知识，并面向不同的推理机制。

基于神经网络的智能诊断专家系统具有传统专家系统所不具有的许多优点，如利用神经网络自身的分布式联结机制对知识进行隐式表示，将知识表示、存储和推理三者融为一体；在知识获取、并行推理、自适应学习、联想推理和容错能力等方面显示出明显的优越性，一定程度上克服了基于符号的传统诊断专家系统存在的知识获取困难，知识存储容量与系统运行速度的矛盾以及知识的窄台阶效应等问题。然而，基于神经网络的智能诊断专家系统也存在固有的弱点。首先，系统性能受到所选择的训练样本集的限制，训练样本集选择不当，特别是在训练样本集很少的情形下，很难指望它具有较好的归纳推理能力；其次，神经网络没有能力解释自己的推理过程和推理依据及其存储知识的意义；再次，神经网络利用知识和表达知识的方式单一，通常的神经网络只能采用数值化的知识；最后，也是最根本的一点是神经网络只能模拟人类感觉层次上的智能活动，在模拟人类复杂层次的思维方面，如基于目标的管理、综合判断与因果分析等方面还远远不及传统的基于符号的专家系统。因而，人们正试图研究符号推理与数值推理相结合的集成式智能诊断系统，以期待能更好地模拟人类的思维过程。

4. 基于案例的诊断

基于案例推理 CBR (Case Based Reasoning) 能通过修订相似问题的成功结果来求解新问题。它通过获取新知识作为案例来进行学习，不需详细的应用领域模型。在 CBR 中的主要技术包括：案例表达和索引、案例的检索、案例的修订、从失败中学习。基于案例的诊断方法的原理是：对于所诊断的对象，根据其特征和症状从案例库中检索出最相似匹配的案例，然后对该案例的诊断结果进行修订作为该对象的诊断结果。基于案例的诊断方法适用于难以表示成规则形式、而易表示成案例形式并且已积累丰富案例的领域（如医学诊断），但它也存在着局限性。传统的基于案例的诊断方

法难以表示案例间的联系；对于大型案例库案例检索十分费时，并且难以决定应选择哪些症状及它们的权重。基于案例的诊断方法难以处理案例修订时的一致性检验（特征变量间的约束关系），难以对诊断结果加以解释。

5. 基于行为的诊断

传统的专家系统是基于知识的人工智能诊断系统，其不足之处在于知识获取困难。而基于行为（Behavior-based）的智能诊断系统是从某一系统的实际运行状态出发，从状态变化判断其状态属性。基于行为的诊断系统是一个动态的、模块化的系统，各个子模块组成相对独立的功能单元。其子诊断模块的数目取决于实际设备的运行行为，随着设备运行时间的增加，出现的病态类别逐渐增多，系统可以自动地添加相应数量的新子诊断模块以对新出现病态的诊断，使其自身的诊断能力得以提高。

基于行为的智能诊断的另一优点是在缺乏先验知识的情况下，仍然能通过与实际行为的交互作用，建立有效的诊断系统。其机制是：即使缺乏病态知识，但正常行为总可以观测搜集到，此时可以用一个子模块识别设备的正常行为，即完成状态监测功能。当设备出现病态行为后，系统自动添加新的子诊断模块，完成对病态的识别，从而逐步增加系统的诊断能力，最终形成完善的智能诊断系统。

基于行为的智能诊断强调依据实际的运行行为建立诊断系统，但并不排斥已经经过实践检验的正确知识和经验，可以将其作为建立系统的重要辅助资源。可以根据已有的知识建立起一个具有初级智能的诊断系统，然后通过与实际行为的交互作用，逐步学习进化为具有高级智能的诊断系统。这样，事先可以不必考虑整个寿命周期中可能遇到的所有病态，从而极大地减轻了病态诊断系统开发的规模和困难。

(二) 诊断技术的发展趋势

1. 集成型诊断系统

诊断与知识表示和推理方法有着密切的联系,其领域知识可用对象模型、经验规则、神经网络模型、案例等来表示。基于对象模型、基于经验规则、基于神经网络模型、基于案例的诊断方法各有其优势和特点,但它们各自也存在着局限性。对于实际对象的病态诊断,如用单一的知识表示方法,有时难以完整地表示对象的病态诊断领域知识,而集成多种知识表示方法则能更好地表示对象的病态诊断领域知识。集成基于对象模型、基于经验规则、基于神经网络模型、基于案例的集成型诊断方法能综合各诊断方法的特点,克服各诊断方法的局限性,从而提高了诊断系统的智能性和诊断效率。集成型的病态诊断系统还能有效地处理真值维护、结论解释、机器学习。

2. 基于机器学习的诊断系统

智能系统的核心问题是它的学习能力。知识的自动获取一直是智能病态诊断专家系统研究中的难点。解决知识获取问题的途径是机器学习,即让机器自身具有获取知识的能力,在实际工作中不断总结成功和失败的经验教训,对知识库中的知识自动进行调整和修改,以丰富和完善系统的知识。机器学习是提高病态诊断系统智能的主要途径,也是衡量一个系统智能程度的主要标志。因此,发展和完善现有的机器学习方法,探索新的学习方法,建立实用的机器学习系统,特别是多种学习方法协同工作的智能诊断系统,将是今后研究的一个重要方向。

3. 由单机诊断到远程分布式全系统诊断

现有的诊断系统大部分都是面向单台、单机或单类设备的,可扩充性、灵活性、通用性较差,各诊断系统之间相互独立,即使是不同开发单位研制的针对同类设备的异构诊断系统之间也不能进行