



Mathematical modeling for active and dynamic diagnosis of crop diseases based on Bayesian networks and incremental learning

Yungang Zhu^{a,b}, Dayou Liu^{a,b,*}, Guifen Chen^c, Haiyang Jia^{a,b}, Helong Yu^c

^a College of Computer Science and Technology, Jilin University, Changchun 130012, PR China

^b Key Laboratory of Symbolic Computation and Knowledge Engineering of the Ministry of Education, Jilin University, Changchun 130012, PR China

^c College of Information Technology, Jilin Agricultural University, Changchun 130118, PR China

ARTICLE INFO

Article history:

Received 30 September 2011

Accepted 26 October 2011

Keywords:

Bayesian networks

Crop diseases

Diagnosis

Incremental learning

ABSTRACT

To achieve rapid and precise diagnosis of crop diseases, an active and dynamic method of diagnosis of crop diseases is needed and such a method is proposed in this paper. This method adopts Bayesian networks to represent the relationships among the symptoms and crop diseases. This method has two main differences from the existing diagnosis methods. First, it does not use all the symptoms in the diagnosis, but purposively selects a subset of symptoms which are the most relevant to diagnosis; the active symptom selection is based on the concept of a Markov blanket in a Bayesian network. Second, a specific incremental learning algorithm for Bayesian networks is also proposed to make the diagnosis model update dynamically over time in order to adapt to temporal changes of environment. Furthermore, the diagnosis results can be calculated without inference in Bayesian networks, so the method has low time complexity. Theoretical analysis and experimental results demonstrate that the proposed method can significantly enhance the performance of crop disease diagnosis.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Diagnosis of crop diseases is an essential task for agricultural production. Crop diseases have become among the most important factors restricting high yield and high quality of crops, and cause major production and economic losses in agriculture. For example, soybean rust (a kind of disease in soybeans) caused yield reduction by 20%, and farmers may have lost approximately 11 million dollars in profit in the year 2006 in the USA [1]. Rice diseases caused yield reduction by 40%–50% at a particularly serious time in China [2]. Therefore, rapid and precise diagnosis of crop diseases to assist agricultural experts is of great significance.

In recent years, many intelligent diagnosis techniques have been developed for crop disease diagnosis. The expert system PLANT/ds developed by the University of Illinois for soybean disease diagnosis was the first agricultural expert system in the world [3]. Tocatlidou et al. [4] proposed an uncertainty reasoning method for crop disease diagnosis by using methodologies and formalisms that provide for representations of facts and rules with various degrees of certainty. Kolhe et al. [5] developed a web-based intelligent disease diagnosis system for diseases of the oilseed crops based on a fuzzy logic approach. Rumpf et al. [6] introduced an approach for detecting sugar beet diseases based on support vector machines. Devraj and Jain [7] designed and developed an expert system for the diagnosis and control of diseases in pulse crops; it can automatically build an integrated and consistent knowledge base combining all the knowledge fed in by the multiple experts and handles uncertainty on the basis of fuzzy logic. Gonzalez-Andujar [8] developed a rule-based method for diagnosis of disease in

* Corresponding author at: College of Computer Science and Technology, Jilin University, Changchun 130012, PR China. Tel.: +86 431 85159333.

E-mail addresses: liudy@jlu.edu.cn, liudyjlu@gmail.com (D. Liu).

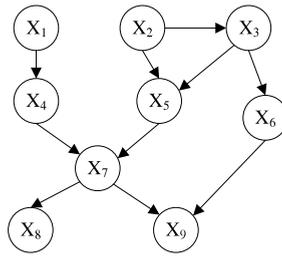


Fig. 1. A simple example of a Bayesian network structure with nine variables.

olive crops. Zhang et al. [9] proposed a multi-disease diagnosis model for tomatoes based on self-organizing map neural networks. Wu and Zhao [10] established a fuzzy matrix for executing C means clustering for disease description, and propose a method of diagnosis of crop diseases based on heuristic and/or graph AO^* search. Sun et al. [11] presented a new knowledge expression method based on weight self-adjustment artificial neural networks for diagnosis of disease in rice. Liu et al. [12] studied the expression method for fuzzy subordination and a model of a fuzzy back-propagation artificial neural network for 26 kinds of common grape diseases, and realized a web-based intelligent grape disease diagnosis system. Sarma et al. [13] described the design and development of a rule-based rice disease diagnosis expert system. Jin and Song [14] proposed an improved decision tree algorithm for crop disease diagnosis.

Although there have been plenty of approaches for crop disease diagnosis, some issues restrict the existing methods. First, the existing methods always use all the disease symptoms for diagnosis. However, because of the noise or imprecise information acquisition devices, not all of the data for symptoms are precise, and not all of the symptoms are the most relevant for diagnosis. Moreover, using more symptoms will have more computation cost, which will affect the diagnosis speed. So this will improve the efficiency of selecting an optimal subset of symptoms which are the most informative and relevant to diagnosis. Second, the existing diagnosis models are always static: after they are established, they are constant. But the environment is always dynamic and changes over time; for example, the most diagnosis-relevant symptoms in summer may be different from those in winter. So the diagnosis model needs to reflect temporal changes of the dynamic environment.

To address the above issues, in this paper we propose an active and dynamic method of diagnosis of crop diseases based on Bayesian networks and incremental learning. In this method, Bayesian networks are adopted for representing crop diseases, symptoms, and the uncertain relationship between them. This method can select a subset of symptoms that are the most relevant to diagnosis purposively (we call this active symptom selection) on the basis of finding a Markov blanket in a Bayesian network. And the diagnosis model can be updated dynamically by incremental learning from newly acquired data, so it has great ability to reflect changes of the environment over time and has great flexibility. Furthermore, the diagnosis result can be obtained without inference in Bayesian networks, so it has low time complexity. The experimental results show that the proposed method can significantly enhance the accuracy and efficiency of crop disease diagnosis.

The remainder of this paper is organized as follows. Section 2 offers brief background knowledge concerning Bayesian networks. The method that we propose is described in Section 3, and Section 4 presents experiments on soybean and cotton diseases using the proposed diagnosis method. Finally, conclusions and recommendations for future work are summarized in Section 5.

2. Theoretical background

A Bayesian network (BN) is a graphical model for representing relationships among variables. Let us consider a set of variables $X = \{X_1, X_2, \dots, X_n\}$; a Bayesian network is a tuple $B = (G, \Theta)$, which includes two components:

- G is a directed acyclic graph (DAG); each node of the graph represents a variable, and each arc represents a direct dependence relationship between variables;
- $\Theta = \{P(X_i|\pi_i), 1 \leq i \leq n\}$ is a set of parameters that represents the conditional probability distribution of each node given the values of their parent nodes, where π_i represents the parent node set of X_i ; Θ is also called the CPT (conditional probability table) for each node.

The joint probability distribution over X can be represented as the product of the CPT of all the variables [15]:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i|\pi_i).$$

For example, Fig. 1 shows a simple Bayesian network structure. In a Bayesian network, the Markov blanket [16] of a node X_i (which can be represented as $MB(X_i)$ for short) is the set of nodes that is composed of its parent nodes, its child nodes and the parent nodes of its child nodes. For example, in Fig. 1 the Markov blanket of variable X_5 is $\{X_2, X_3, X_4, X_7\}$.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات