

# QPAIS: A WEB-BASED EXPERT SYSTEM FOR ASSISTED IDENTIFICATION OF QUARANTINE STORED INSECT PESTS

Han Huang<sup>1</sup>, Edwin G. Rajotte<sup>2</sup>, Zhihong Li<sup>1,\*</sup>, Ke Chen<sup>3</sup>, Shengfang Zhang<sup>3</sup>

<sup>1</sup> *Department of Plant Protection and Quarantine, China Agricultural University, Beijing, P. R. China 100193*

<sup>2</sup> *Department of Entomology, Penn State University, State College, Pennsylvania, USA 16802*

<sup>3</sup> *Animal and Plant Quarantine Institute, China National Academy of Inspection and Quarantine, Beijing, P. R. China 100029*

\* *Corresponding author, Address: Department of Plant Protection and Quarantine, College of Agronomy and Biotechnology, China Agricultural University, 100193, Beijing, P. R. China, Tel: +86-10-62731299, Fax: +86-10-62733404, Email: lizh@cau.edu.cn*

**Abstract:** Stored insect pests can seriously depredate stored products causing worldwide economic losses. Pests enter countries traveling with transported goods. Inspection and Quarantine activities are essential to prevent the invasion and spread of pests. Identification of quarantine stored insect pests is an important component of the China's Inspection and Quarantine procedure, and it is necessary not only to identify whether the species captured is an invasive species, but determine control procedures for stored insect pests. With the development of information technologies, many expert systems that aid in the identification of agricultural pests have been developed. Expert systems for the identification of quarantine stored insect pests are rare and are mainly developed for stand-alone PCs. This paper describes the development of a web-based expert system for identification of quarantine stored insect pests as part of the China 11th Five-Year National Scientific and Technological Support Project (115 Project). Based on user needs, textual knowledge and images were gathered from the literature and expert interviews. ASP.NET, C# and SQL language were used to program the system. Improvement of identification efficiency and flexibility was achieved using a new inference method called characteristic-select-based spatial distance method. The expert system can assist identifying 150 species of quarantine stored insect pests and provide detailed information for each species. The expert system has also been evaluated using two steps: system testing and identification testing. With a

85% rate of correct identification and high efficiency, the system evaluation shows that this expert system can be used in identification work of quarantine stored insect pests.

Keywords: expert systems, quarantine stored insect pests, identification, stored products

## 1. INTRODUCTION

Stored products play an important role in every country because they ensure a food supply and strengthen the stability of national economies. Generally, stored products contain grain, oil, medicinal materials, tobacco, fur, books, file documents, etc. They are usually preserved in environments which include storehouses, workshops, warehouses, libraries, etc. Stored pests are those that live in these environments and include insect pests, mites, rodents, etc. Stored pests have the ability to adapt themselves to stored product environments and can cause severe damages to every kind of stored product by feeding and contaminating them. Research from USDA (United States Department of Agriculture) showed that stored pests caused 2.3 billion dollars in losses during the storing and processing of agricultural products in 1960s (Yin et al., 2002). In China, during the annual storage period in 1980s, the loss caused by stored pests in grain oil was about 10%, losses in Chinese traditional medicine materials were about 70% (Li et al., 2000). In order to reduce losses against stored pests, about 3 billion kilograms pesticides were used worldwide, which contributes to environmental pollution (Pinentel, 1991). Stored pests are now being widely researched to better manage them across the world because of their astounding ability to destroy products and their economic importance. Although some success has been attained in recent years using a variety of preventative methods including physical prevention, chemical prevention and biological prevention, it is still a long-term and complicated task to control stored pests, especially preventing the spread of stored pests from one country/place to other countries/places. With more and more international trade and communication, it becomes increasingly likely that stored pests are able to move with goods. Inspection and Quarantine work is a vital procedure for every country in order to control the spread of domestic stored pests and to prevent the invasion of foreign stored pests, especially quarantine stored pests. There are 533 stored pest species known in the world and 383 of these exist in China, stored insect pests account for the majority of them (Yin et al., 2002). In China, only a few parts of stored insect pests are quarantine stored insect pests (Xu, 1994), and are mainly in

the families *Dermestidae* and *Bruchidae*. Identification is an important step in the process of Inspection and Quarantine work to identify species captured. Traditionally in China, species determination by personnel trained in insect identification was the norm, but the requirements of Inspection and Quarantine work under the new international era, the lack of taxonomic specialists, low efficiency, and other factors necessitate other approaches. The rapid development and the wide use of information technologies in different fields present the opportunity to make insect identification more efficient.

An expert system, a kind of intelligent computer program, uses knowledge and inference to solve problems that are difficult enough to require human domain experts. Expert systems have been developed for identification applications in every field of agriculture, involving insect pests, weeds, poultry diseases, etc. (e.g. Liu et al., 2002; Tang, 1999; Feng et al., 2007; Travis et al., 1992). In the identification of stored insect pests, only a few expert systems have been reported (e.g. Li et al., 2006; Li et al., 2001). Li Kaibing et al. (2006) developed a *Scolytidae* Identification System, which is a multimedia expert system for identification of *Scolytidae* stored insect pests involved in China Inspection and Quarantine work, this expert system is based on Taxokeys that is an insect classification platform (Gao, 2003).

All of the existing expert systems for identification of stored insect pests only refer to a few taxa within the quarantine stored insect pests and were developed for stand-alone PCs. The inference method of identification used in these expert systems is the traditional key-based binary tree search method which takes users much time and is less flexible. With the rapid development of Internet technologies, building a web-based expert system with a more flexible and highly efficient inference method for identification of all quarantine stored insect pests is essential to China's inspection and quarantine work.

A web-based expert system for identification of quarantine stored insect pests, called QPAIS (Quarantine Pests of Stored Products Assistant Identification System), has been developed by the Plant Quarantine Laboratory at China Agricultural University. It is an outcome of the China 11th Five-Year National Scientific and Technological Support Project (115 Project). This web-based intelligent system contains a large amount of detailed textual information and images of about 150 quarantine stored insect pest species. An array of inference knowledge with a new inference method for identification was also included in this system. The inference method is a characteristic-select-based spatial distance method with high efficiency and high accuracy. The expert system has been tested and will be used in China's Inspection and Quarantine work. In addition, the system can also be used in educational training.

## 2. DOMAIN BACKGROUND

### 2.1 Brief introduction of taxonomy

Quarantine stored insect pests in China mainly refer to species in the *Dermestidae* and *Bruchidae*, which belong to the beetles (*Coleoptera*). These species have a common life cycle, which consists of four life stages including egg, larva, pupa and adult. The adults of these beetles have small, ovoid bodies covered with scales or hairs of different colors. The main parts of the adult body include the head, thorax and abdomen. Also common to the adult are a pair of antennae of varying shapes, three pairs of legs, two pairs of wings, with first pair hardened into elytra.

### 2.2 General process of identification

Compared with the other three life stages, the adult is easier to identify because of distinctive characteristics, such as the shape of antennae, different color hair zones on the body and elytra, the shape of pronotum, etc. Species are classified into a number of groups based on common characteristics, all groups have a hierarchical relationship with each other. For example, the species *Dermestes dimidiatus* belongs to the genus *Dermestes* which belongs to the family *Dermestidae*. There must be at least one different characteristic between any two species which can be used in identification.

In the traditional identification process of quarantine stored insect pests, the first step is to observe all the characteristics of an unknown adult, usually by eyes or with a microscope. The second step is to determine the species by comparing all the observations with knowledge including text and images from books or identification experience. Sometimes, these two steps are simultaneous.

## 3. USERS' NEEDS

According to interception records of quarantine stored insect pests from China inspection offices at ports during the past decade, transportation of goods is the major avenue for pests to spread with from one country to another. Identification work is vital to determine whether a species captured is a foreign invasive species or a domestic species, and then help to make a decision about the disposal of the goods containing pests.

In order to meet the requirements of China's Inspection and Quarantine work under the new international era, an inquiry was conducted by China National Academy of Inspection and Quarantine (CAIQ) to find out the practical needs of China's inspection offices at the ports. The results show that these offices have many needs:

1. Lack of classification experts. The identification work can be accomplished by classification experts who possess the proper knowledge and abundant practical identification experience. China is a large country and has many port inspection offices. People working in these offices have only basic knowledge of taxonomy and still need to consult taxonomy experts when a species which is not familiar to them is captured. As the number of taxonomy experts in China is limited, the identification shortfalls at inspection offices cannot be resolved.

2. Long time needed. If a classification expert is not available, identification work often requires inexperienced port personnel to spend much time consulting literature in order to make identifications.

3. Lack of systematic data. Detailed information, such as taxonomic position, quarantine status, classification characteristics, biology characteristics, geographic distributions, etc., is difficult to acquire. People working in port inspection offices still need to gather the information from many reference books or the internet.

A web-based assistant identification system, which can be used in China's Inspection and Quarantine offices will help resolve the problems mentioned above.

#### **4. KNOWLEDGE ACQUISITION**

A critical aspect of building an expert system is formulating the scope of the problem and gleaning from the source expert the domain information needed to solve the problem (Gonzalez-Andujar, 2006). This aspect can be achieved through knowledge acquisition, which is an important process and has always been considered a bottleneck during the development period of expert systems. This bottleneck is mainly caused by communication difficulties between the knowledge engineer (KE) and the domain expert, the inability of the expert to describe his/her expertise, and the inability of the KE to understand and codify the expertise (Liebowitz et al., 1996; Li et al., 2002). In the present system, two approaches have been adopted to achieve knowledge acquisition.

## 4.1 Literature reference

Textual information (e.g. species' morphological characteristics, biology, geographic distribution, etc.) and images of 150 species of quarantine stored insect pests in *Dermestidae* and *Bruchuidae* were acquired from literature such as textbooks, primary and secondary literature, papers, etc. We also collected some images from the China National Academy of Inspection and Quarantine (CAIQ). All these quarantine stored insect pest species, which account for bulk of stored pests, were commonly found in international trade all around the world.

## 4.2 Experts interviews

Through years of experience, taxonomy experts have developed a body of knowledge and classification experience which they can use to make correct identifications (Li et al., 2002). The famous taxonomy expert of stored insect pests from China National Academy of Inspection and Quarantine (CAIQ) took part in knowledge acquisition.

Table 1. *Dermestes* species considered in QPAIS with Abdomen hair Identification characteristic and its choices

Dermestes species	Identification characteristics		
	Abdomen hair		
Latin name	Having thickly white or light yellow hairs which cover abdomen	Having thinly light yellow or dark color hairs which don't cover abdomen	Not clear
<i>Dermestes dimidiatus</i>	×		×
<i>Dermestes elegans</i>	×		×
<i>Dermestes marmoratus</i>	×		×
<i>Dermestes coronatus</i>	×		×
<i>Dermestes maculates</i>	×		×
<i>Dermestes carnivorus</i>	×		×
<i>Dermestes frischii</i>	×		×
<i>Dermestes coarctatus</i>	×		×
<i>Dermestes sibiricus</i>	×		×
<i>Dermestes lanarius</i>	×		×
<i>Dermestes murinus</i>	×		×
<i>Dermestes undulates</i>	×		×
<i>Dermestes tessellatocollis</i>	×		×
<i>Dermestes lardarius</i>		×	×
<i>Dermestes vorax</i>		×	×
<i>Dermestes ater</i>		×	×
<i>Dermestes leechi</i>		×	×
<i>Dermestes freudei</i>		×	×
<i>Dermestes nidum</i>		×	×
<i>Dermestes haemorrhoidalis</i>		×	×
<i>Dermestes peruvianus</i>		×	×

The taxonomy expert picked related identification characteristics most frequently used in identification work in different groups (e.g. *Coleoptera*, *Dermestidae*, *Bruchidae*, etc.), each identification characteristic includes some characteristic choices. The taxonomy expert also added a ‘not clear’ option to every characteristic as the final choice to accommodate absent or unclear characteristics of samples. For instance, Table1. shows 16 characteristics, including abdomen hair, vaginal hair color, vaginal hair length, etc, for identifying species in *Dermestes* which contains 21 *Dermestes* species. The abdominal hair characteristic has 3 characteristic choices, two of which can be chosen to classify 21 *Dermestes* species into two smaller groups, the third one is ‘not clear’ which has no classifying function. Also, each characteristic choice was given a mark value which was defined as the degree of relatedness between characteristics and specific species. Mark values can range from 0-100 and are based on expert opinions. Table2. contains an example of three characteristic choices and related mark values of the scutellum. ‘Mahogany hair on the scutellum’ is the characteristic only to *Dermestes lanarius* but not to others in *Dermestes*, so the mark value of mahogany hair on the scutellum is 100.

Table 2. Characteristic choices of scutellum with mark values in *Dermestes*

Identification characters	
Scutellum	Mark Value
Characteristic choices	
Mahogany hair on the scutellum	100
Yellow hairs on the scutellum	100
Not clear	0

## 5. QPAIS EXPERT SYSTEM DESIGN AND DEVELOPMENT

QPAIS expert system consists of a user information database, a stored insect pest information database, a knowledge database, an inference engine and a user interface. The structure is represented in Fig.1.

### 5.1 Development software selection

QPAIS was designed to run on the internet. Its development was based on the use of internet techniques and SQL programming languages. HTML (Hypertext Markup Language), Java Script, C# and ASP.NET were used in the programming. Other software included MS SQL Server 2000, Windows XP, IIS 6.0, Photoshop 7.0 and Flash 7.0.

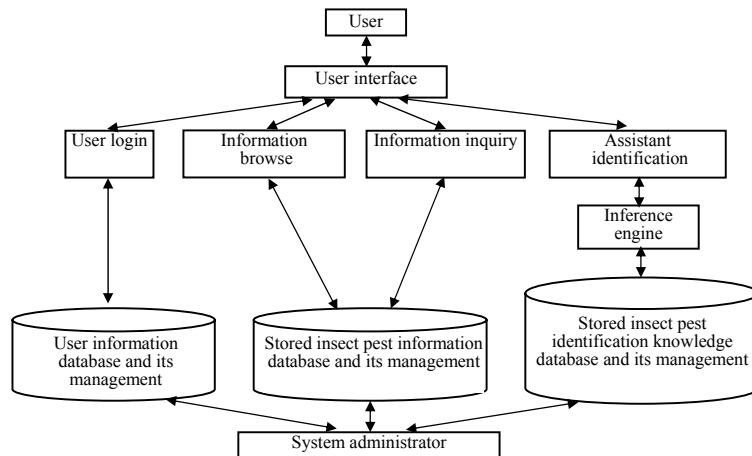


Fig. 1: System structure of QPAIS

## 5.2 Databases

The server database played a very important role in developing QPAIS. It was used to sort all the information and knowledge which was needed to actualize every function of the expert system. In QPAIS, there are three databases including a user information database, a basic information database and an identification knowledge database. They were all designed using MS SQL Server 2000.

The user information database contains the users' name, password, rights, etc. This makes it easier for system administrators to manage users and authorize the access to different functions. The basic information database includes detailed textual information of quarantine stored insect pests, including Chinese names, scientific names, taxonomic position, quarantine status, figure descriptions, biological characteristics, geographic distributions, etc. Images related to species were also included in the basic information database. In the identification knowledge database, identification characteristics, characteristic descriptions and related mark values gathered from taxonomy experts were stored.

## 5.3 Inference process

The SQL language was used to operate the identification knowledge database and implement the inference process in QPAIS. The inference



process uses characteristic-select-based spatial distance method and is shown in Fig.2.

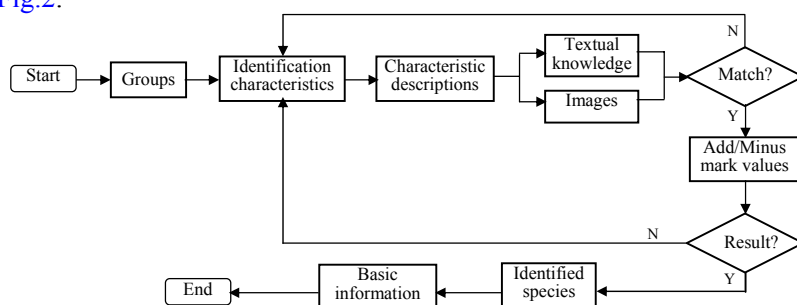


Fig.2: QPAIS inference process

For example, for an unknown species, the process starts by determining which group this species should belong to. If this species can be confirmed as a member of *Dermestes*, the system will display 16 identification characteristics. Each characteristic can be chosen to show related characteristic choices. When choosing the ‘scutellum’ characteristic, the system will show three characteristic choices with related images. If the ‘Mahogany hair on the scutellum’ choice, which is the key characteristic only to *Dermestes lanarius*, is chosen, *Dermestes lanarius* will be added to the result list with a mark value of +100, each of the other 20 species in *Dermestes* will get a -100 mark value because they do not have that characteristic. If the ‘Abdominal hair’ characteristic is chosen as well, the system will list three corresponding characteristic choices. If ‘Having thickly white or light yellow hairs which cover abdomen’ is chosen, 13 species including *Dermestes lanarius* will be assigned +20 mark values while the other 8 species will be assigned -20. After two selections, *Dermestes lanarius* has a total mark value of +120 (100+20), 12 species having ‘Having thickly white or light yellow hairs which cover abdomen’ characteristic respectively get -80 mark values (-100+20), other 8 species respectively have -120 mark values (-100-20) as their final score. The final result list will show species having scores bigger than 0 from high to low, obviously, *Dermestes lanarius* with 120 mark values is the only one species in the result list. Then *Dermestes lanarius* can be regarded as the name of the species to be identified.

If the species cannot be confirmed as the member of a specific genus, the system will provide family identification characteristics and related characteristic choices. The result list with mark values will show *Dermestidae* or *Bruchidae* as the top item that this species should belong to. After this step, the species can be identified as the member of certain genus

in *Dermestidae* or *Bruchuidae* by choosing related genus identification characteristics.

### 5.4 Users interface

Users find it easy to operate the multimedia interface. There are abundant combinations of textual information and images in the interfaces. Users have three main options: information browsing, information inquiry, identification assistant. On the ‘browsing information’ page, the system displays a list of quarantine stored insect pests and the user can inspect detailed information related to a specific species after choosing its ‘browse’ link. If the user clicks the option ‘information inquiry’, the system provides an array of inquiry textboxes for users to input keywords to inquire about a specific species.

As shown in Fig.3, the ‘identification assistant’ page firstly offers a number of groups (e.g. *Dermestidae*, *Bruchuidae*, etc.) for selecting. If one of the groups is selected, the system will list related identification characteristics and characteristic choices for users to choose, all characteristic choices are combined with thumbnail sized images which can be expanded to larger ones by clicking. Fig.4 shows the identification characteristics and characteristic choices. Finally, the system shows the pest species at the top of a list which contains possible quarantine stored insect pests with different scores from high to low. Users can further check the detailed information of the identified species by clicking on linkbutton ‘browse’, which is displayed in Fig.5. The system has a record of every step in the identification process and allows users to jump back to previous steps to make another choice.



Fig.3: A screen shot for user interface on group selection



Fig.4: An example of an identification interface used to identify a species belonging to *Dermestes* with thumbnail sized images which can be expanded to larger ones by clicking



Fig.5: An example of an identification outcome interface with link button which can be clicked on to browse the detailed information of the identified species

## **6. SYSTEM EVALUATION**

System evaluation is a process to ensure expert systems can work correctly and meet the requirements as designed. System evaluation of QPAIS contains two steps: the system testing and the identification testing. In the system testing step, function debugging, logic testing and rule checking were carried out by system developers as well as testing participants who include two programmers from China National Academy of Inspection and Quarantine (CAIQ) and 15 graduate students from Plant Quarantine Laboratory at China Agricultural University. Every function was tested for hundreds of times until it was confirmed to have no errors. Every error found was recorded for system developers to further debug.

The identification testing is a complex step which needs the taxonomic knowledge and identification experience of taxonomy experts. The taxonomy expert from China National Academy of Inspection and Quarantine (CAIQ) was responsible for this step work. A total of 150 quarantine stored insect pests were used to test the correctness of identification results. After debugging of characteristics, choices and related mark values specified by the taxonomy expert, QPAIS identification system was able to identify 85% of the cases correctly. The erroneous identifications were the consequence of misidentifications of a species and some similar species.

With the information gathered from the system evaluation, QPAIS was further improved.

## **7. CONCLUSIONS AND FUTURE DEVELOPMENTS**

The aim of developing QPAIS was to help identify quarantine stored insect pests with high efficiency in China's Inspection and Quarantine work, as well as provided educational value. The taxonomy experts and students who used this system to identify certain species had a common opinion that QPAIS really helps improve the identifications of quarantine stored insect pests with high accuracies, and it can be used in the port inspection offices.

This system contains detailed information about 150 stored insect pest species that are commonly found in international trade and stored environments. With a large amount of textual knowledge and related images, it is easy to identify species. QPAIS is different from traditional key-based binary tree methods used in identification expert systems. QPAIS uses a characteristic-select-based method, which allows users to get the identification result in a short time just by selecting several identification

characteristic choices. QPAIS was developed to run on the internet, which is convenient for users throughout the world

Based on the records from the system evaluations, further developments are under considerations in order to make the system more helpful. This system is a prototype system and contains 150 quarantine stored insect pest species at present, to better meet the requirements of quarantine and management work, more species should be added. With the rapid development of Artificial Neural Network (ANN) in every kind of field, we are planning to refine the identification method toward ANN in the future, though it is similar to the identification method already used in QPAIS.

## ACKNOWLEDGEMENTS

We would like to thank Wenxin Li and Zhigang Wu for helping resolve some difficulties in system requirement analysis and programming. Thanks also should go to graduate students in Plant Quarantine Laboratory at China Agricultural University for their testing work. The funding of this study was supported by China 11th Five-Year National Scientific and Technological Support Project 'Port inspection technologies of potentially invasive species' (2006BAD08A13).

## REFERENCES

- David Pinetel. World resources and food losses to pests, *Ecology and management of food industry pests*, 1991, 5-11
- Feng Wanyu, Xiao Jianhua, Liu Yun, et al. Studying the diagnostic expert system of swine diseases on based of integrated technology, *Journal of stock and veterinarian*, 2007, 26(2):27-28(in Chinese)
- Gao Lingwang, Shen Zuorui, Liu Zhiqi. Design and development of Taxokeys: a dichotomous-reasoning-based multimedia expert system assisting insect identification and taxonomic study, *Entomology*, 2003, 46(5):644-648(in Chinese)
- J. L. Gonzalez-Andujar, C. Fernandez-Quintanilla, J. Izquierdo, et al. SIMCE:An expert system for seedling weed identification in cereals, *Computers and electronics in agriculture*, 2006, 54:115-123
- J. Liebowitz, S. I. Baek. The protocol multimedia expert system, *The New Review of Applied Expert Systems*, 1996, 1:3-17
- J. W. Travis, E. G. Rajotte, R. Bankert, et al. Penn State apple orchard consultant expert system: The design and function of the pest management module, *Plant Disease*, 1992, 76(6):545-554
- Li Daoliang, Fu Zetian, Duan Yanqing. Fish-Expert: a web-based expert system for fish disease diagnosis, *Expert Systems with Applications*, 2002, 23:311-320

- Li Kaibing, Gao Lingwang, Shen Zuorui, et al. Development of the expert system for assistant identification of insects of scolytidae based on the platform of Toxakeys, Plant quarantine, 2006, 20:17-19(in Chinese)
- Li Shulong, Zhao Zhimo. Review and prospect of research and control of insects in storehouse in China, Entomological Knowledge, 2000, 37(2):84-88 (in Chinese)
- Li Zhihong, Zhang Baofeng, Chen Hongjun. Expert systems and assistant identifications of quarantine pests, Plant quarantine, 2001, 15(4):235-239(in Chinese)
- Liu Yuexian, Shen Zuorui, Cai Xinyan. Research of Agricultural pests assistant identification and control consultation system, Computers and Agriculture, 2002, 1:9-11(in Chinese)
- Tang Yuechen, Chen Jianwu. Weeds identification expert system, Journal of Fujian Agricultural University, 1999, 28(3):330-334(in Chinese)
- Xu Guogan. Stored pests quarantine, Grain storage, 1994, 23:100-104(in Chinese)
- Yin Wenya, Wang Xiaoping, Zhou Chengai. Pests in storehouse and status of chemical control, Hunan Agricultural Science, 2002, 6:54-56 (in Chinese)