PBIS – Computer-user Interface for Phenotypic Breed Identification using Scoring-function Methodology

Avnish. K. Bhatia[1],#, Pradeep Gupta[1], Aanand. Jain[1], D. K. Sadana[1], S. B. Gokhale[2], R. L. Bhagat[2]

National Bureau of Animal Genetic Resources, Karnal – 132001

ABSTRACT

This paper presents computer user-interface, viz., Phenotypic Breed Identification System (PBIS) developed for scoring-function methodology devised for animal's breed identification based on phenotypic traits (Bhatia et al. 2010, Computers & Electronics in Agriculture, 73: 37-43). PBIS has facilities for adding basic information on livestock species, species-wise animal characteristics, breeds and breed descriptors. Users can add and view test animals along with their phenotypic traits. Breed of test animals can be identified by selecting animals, list of traits and list of breeds for comparison. Results of identification are presented as a report containing list of animals, user's assessment of breed conformation in percentage, system identified breed, list of breeds compared and list of traits.

Keywords: User-interface, Animal genetic resources, Breed, Computer learning

Present Adresses:[1] National Bureau of Animal Genetic Resources, Karnal – 132001, [2]. BAIF Central Research Station, Uruli Kanchan, Pune – 412202; # Email: avnish@lycos.com

INTRODUCTION

Animal genetic resources of a region include different breeds of domesticated animal species such as cattle, buffalo, sheep, goat, camel etc. being reared in the region. These breeds are distinguished on the basis of certain morphological and production traits. Identification of breeds of animals is an important task for improvement and conservation of indigenous animal genetic resources. It is also required for breed-wise livestock census as well as for animal trade across borders.

Basic information on a breed is available in the form of breed descriptors separately for each category of animals such as adult females, adult males, young-ones, draft animals, etc., (Bhatia, Yadav and Vij 2009). The breed descriptors consist of mean values along with standard error for quantitative traits and frequencies of trait values in case of qualitative traits, and are available in field survey reports and information systems on animal genetic resources (Ayalew 2003). Breed descriptors for a number of indigenous breeds of livestock and poultry species are available in India (Joshi and Vij 2009).

Animals are generally identified to a breed visually with human judgment on the basis of phenotypic traits. The problem of identification of breed of animals can be solved with computer learning under the purview of classification of unknown samples. Classification aims to predict the unknown output value / class for a given input vector of feature values on the basis of a set of pre-classified samples. Artificial neural networks, decision trees, instance-based learning and Bayesian classification are some of the popular computer learning algorithms (Huang et al. 2010). All the learning algorithms require a set of pre-classified samples for decision making on the class of test samples. As such, these can't be used for identification of breed of animals on the basis of available breed descriptors consisting of analysed information.

Recently, a computer learning methodology based on scoring-function has been developed for making objective decision on identification of breed of animal using phenotypic traits of the animal and available breed descriptors (Bhatia et al. 2010). It utilizes breed descriptors to calculate scores of a test animal for all the breeds under comparison and identifies the animal to the breed with maximum score. The scoring-based methodology has compared well with existing instance-based learning methods on both simulated and actual data on animals (Bhatia et al. 2010).

Users interact with modern computers through user-interfaces that provide a level of support to utilize systems and methodologies of data analysis and decision making (Abbas and Aggarwal 2010). A computer user-interface, viz., Phenotypic Breed Identification System (PBIS) has been developed for the scoring-function methodology for making decision on the breed of an animal. The remaining paper describes development and use of PBIS.

MATERIAL AND METHODS

2.1 Scoring-function Methodology

Scoring-function methodology has been introduced in Bhatia et al. (2010). It uses a scoring function (1) to calculate score of an animal for a breed using trait values and weights. The scoring function generates a score of the animal for a breed on the basis of respective breed descriptor.

$$S = \sum_{i=1}^{n} w_i v_i \qquad (1),$$

where S is the calculated score of an animal for a breed, n is the number of traits utilized, w_i is the weight assigned to i^{th} trait, and v_i is the value of the i^{th} trait.

A maximum score equal to one hundred has been taken. Sum of all the weights has been taken equal to the number of traits. Maximum value of i^{th} trait has been taken as: $v_i^{max} = 100/n$

In case of qualitative traits, $v_i = v_i^{max} \cdot F_i$ where F_i is the frequency of the value of trait of the animal in the breed descriptor. In case of quantitative traits, $v_i = v_i^{max} \cdot (1 - D_i)$ where $(1 - D_i)$ is the distance of the value of i^{th} trait of the animal from the mean value of the trait in the breed descriptor. $D_i = CSND(z_i) - CSND(-z_i)$, $CSND(z)$ is cumulative standard normal distribution value of z, z is the standard normal variate.

The scoring function provides a score with maximum limit of one hundred for any number of traits and distribution of a total weight among the traits. Scores for an animal is calculated for each breed under comparison. The breed with the highest score among the calculated scores is identified as breed of the animal. Decision on an animal as belonging to a single breed under comparison is made on the basis of location of score of the animal among the calculated scores of one thousand animals generated from breed descriptor. The procedure is also used to make a decision when the animal does not belong to any of the breeds under comparison (Bhatia *et al.* 2010).

2.2 Computer Resources

PBIS uses a database that has been designed to store users, breed descriptors, test animals and result of the scoring function methodology. The database includes tables for storing users, species, breeds, species-wise traits, breed descriptors for different categories of animals, test animals along with their trait values, and results of breed identification. The database in the system has been created using MS-Access 2003. Computer program of the system has been developed using Visual Basic programming language in Microsoft Visual Studio 2008 environment. Crystal report available in MS Visual Studio has been utilized for final report generation.

2.3 Available Breed Descriptors

Available breed descriptors of breeds of cattle, buffalo, sheep and goat species have been entered into the database from respective detail survey reports under Network Project on Animal Genetic Resources (NBAGR 2005, NBAGR 2008). Table-1 shows the list of breed descriptors along with traits of livestock species contained in the PBIS. The table displays list of qualitative traits having frequency of trait values, and quantitative traits having mean trait value along with standard error and number of observations.

Table-1: Traits and breed descriptors entered into PBIS from detailed survey reports

Species	Breeds	Quantitative traits	Qualitative traits
Cattle	Dangi, Khillar, Deoni, Gaolao, Gir, Gangatiri	Body length (cm), Chest girth (cm), Ear length (cm), Head length (cm), Height at withers (cm), Horn size (cm)	Body size, Coat color, Dewlap, Horn color, Horn orientation, Horn shape, Hump size, Muzzle color, Naval flap, Poll, Tail switch color, Teat shape, Teat tip, Udder shape
Buffalo	Jaffarabadi, Nagpuri, Surti	Body length (cm), Chest girth (cm), Ear length (cm), Head length (cm), Height at withers (cm), Horn size (cm)	Body size, Coat color, Dewlap, Horn color, Horn orientation, Horn shape, Hump size, Muzzle color, Naval flap, Poll, Tail switch color, Teat shape, Teat tip, Udder shape
Goat	Attapaddi, Mehsana, Osmanabadi	Adult Body weight (kg), Body length (cm), Height at withers (cm), Chest-girth (cm)	Body size, Coat type, Coat colour, Face colour, Wool cover (Head, Face, Belly, Legs), Head profile, Ear orientation, Ear type Ear length, Horn size Horn shape, Staple Length Fibre diameter, Lustre, Crimp
Sheep	Bonpala, Chhotanagpuri, Coimbatore, Mandya, Rampur bushair, Changthangi, Deccani	Adult Body weight (kg), Body length (cm), Height at withers (cm), Chest-girth (cm)	Body size, Coat type, Coat colour, Distinctive colour markings, Beard, Head profile, Ear orientation Ear type, Ear length Horn shape, Horn orientation, Horn size, Wattles, Average daily milk yield

DESCRIPTION OF PBIS

Users of the PBIS have been categorized into two groups – Organisers and Ordinary users. Figure-1 shows program flowchart for both the categories of users. Organisers can manage the basic information on breeds such as addition of new livestock species, species-wise list of traits, trait options for qualitative traits, breeds and breed descriptors. Ordinary users can enter the test animals and their phenotypic trait values. Conformation of an animal to a breed is also entered according to user's perception. Breed of test animals can be identified for all the breeds with breed descriptors available in the database.

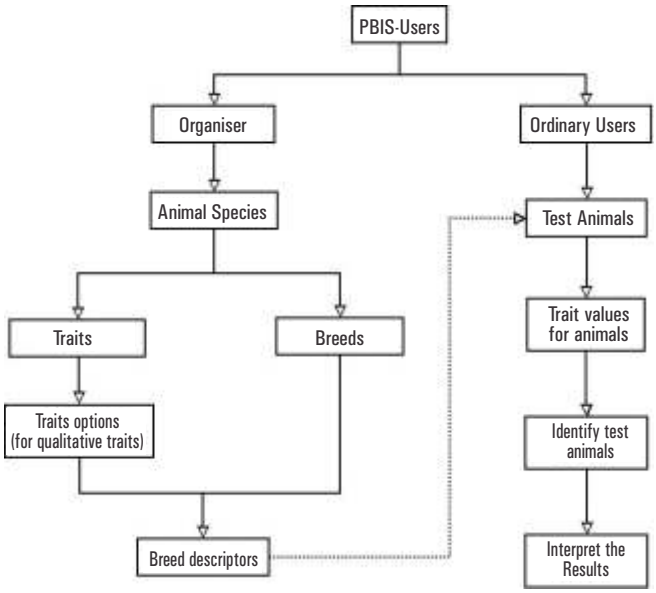

Figure 1: PBIS flowchart for Organiser and Ordinary Users

User Interface Forms: A user login form is displayed on execution of the PBIS that accepts login and password from the user. If the login name does not exist, the user is required to create a new user-login. PBIS Main Menu contains options for ordinary user and the organiser.

Facilities for Organisers: 'Organiser' menu option has sub-menu options for addition of species, breed, traits for a species, breed descriptor and breed-wise weights. The sub-menu item 'Add Animal Species' facilitates addition of a new livestock / poultry species if it is not available in the database. 'Add New Breed' opens form for adding new breed for a species. The sub-menu 'Add Traits for a Species' has two sub-menu options: "Add Traits" and "Add Trait Options". 'Add Traits' opens a form to facilitate addition of species-wise traits on which the breed descriptors of breeds are defined. 'Add Trait Options' opens a form to facilitate addition of trait options. If the trait is qualitative that accepts text type of data, there are options for the trait values. For example, black, white, brown, etc. are options for a trait 'body color'.

Sub-menu 'Breed Descriptor' facilitates addition and viewing of the trait values for a breed through 'Add' and 'View'. 'Add Breed Descriptor' form facilitates entry of trait values for a breed. User selects a trait and enters its value for the selected breed. The selected trait may have numeric value along with standard deviation or standard error with number of observations. If standard error is available, entry of number of observations is essential to calculate standard deviation at the time of computation. In case of qualitative traits, users select an option and then enter frequency value (in percent) of the option for the breed. 'View Breed Descriptor' form facilitates views on the breed descriptors of various breeds.

Sub-menu 'Breed-wise Weights' facilitates an organiser to assign higher weights to a few traits for a breed, if the breed is exclusively identified with these traits. By default, each trait has a weight equal to one. If higher weights are assigned to a few traits, weights are proportionately distributed among all the traits so as to get the sum of all the weights equal to the number of traits. *Note:* Performance of the scoring function methodology with assignment of higher weights to a few traits has not been experimentally verified. It is for the user to assess benefit of assigning higher weights to traits.

Facilities for Ordinary Users: This category of users can view 'User' menu that has three sub-menus for adding, viewing and identifying test animals. 'Add Test Animals' has two menu options: 'Animals' and 'Traits of Animals' facilitating addition of test animals and entry of trait values for test animals respectively. In the form 'Add Test Animals', user enters breed of the animals and breed conformation in percent according to his perception. The form 'Animal Traits' facilitates entry of traits of test animals. If trait value is numeric, it accepts a number. If trait value is of type text, it accepts the trait value as a trait-option from the list. Users can view the animals along with their trait values by opening the form 'View Test Animals'.

Form to identify breed of animals is displayed by selection of menu item 'Identify Test Animals'. Figure-2 shows the form to identify breeds. It displays list of test animals after selecting species, breed and category of animals. Breed names with available breed descriptors are listed. User can select one or more items from the lists. All the items can also be selected by checking the respective check-buttons.

'Simulation Size' is useful in decision support on breed of animals. It generates hypothetical animals equal to simulation size from the breed descriptors of selected breeds and calculates scores of these animals. It displays (in brackets) percent of simulated animals with scores below the score of a test animal for a breed in the final report. It helps a user to make decision about conformity of an animal to a breed. If scores of most of the simulated animals is below the score of a test animal, its conformation to a breed is high. Usually simulation size should be equal to 1000. Simulation of animals is also useful in cases when (i) the test animals do not belong to any of the breeds selected in the list. (ii) Only one breed is selected from the list (Bhatia et al. 2010).

Check-box 'Use Weights' tells the system whether to use weights assigned to traits for a breed or not. The system assigns a weight equal to one to each trait if the box is not checked. Finally, 'Identify Animals' starts calculation of scores to identify breed of test animals using the scoring-function methodology.

Display of Results: Figure-3 shows the final report as result of breed identification. It displays allotted number of animal along with its

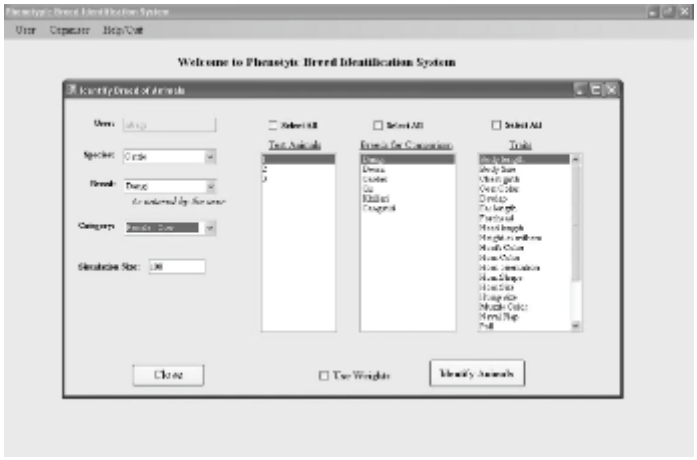

Figure 2: Form to identify animals to a breed

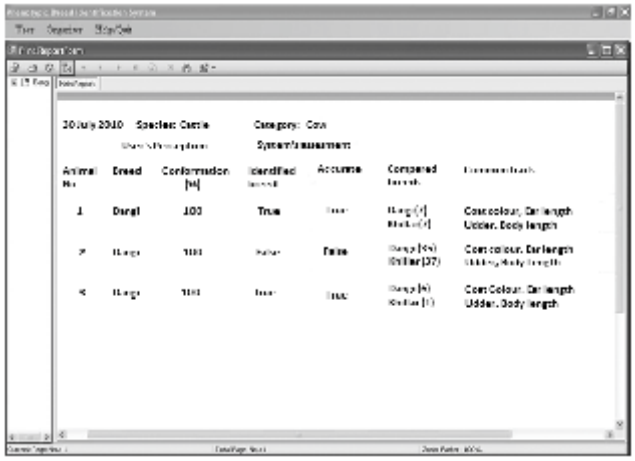

Figure 3: Result of phenotypic breed identification

breed and conformation according to user's perception. Alongside it also displays the system's assessment according to scoring-function methodology as system-identified breed and its accuracy value as 'true' or 'false'. List of breeds compared and traits utilized are also displayed. A value is displayed along with each of the compared breeds, which is percent of simulated animals with score for a breed less than the score of test animals for the breed. This value is taken for decision support on breed of the animal.

Availability and Use: PBIS is a stand-alone package available on a compact disc medium and can be obtained on request. It can be installed on a computer through a setup file. It also includes a MS-Word file for user's guidelines and a help subsystem with a link in the menu. PBIS requires Microsoft .NET Framework 3.5 should be installed on the computer. System requirements of PBIS are same as that of Microsoft .NET Framework 3.5, which runs in Windows Server 2003, Windows Server 2008, Windows Vista and Windows XP and later versions of Windows. It requires 400 MHz Pentium processor (Minimum), 96 MB RAM (Minimum), up to 500 MB hard disk space, and 800 x 600, 256 colours display (Minimum).

REFERENCES

Abbas SQ and Aggarwal A. 2010. Development of a quality design framework for usable user interfaces. *International Journal of Computer Science and Engineering* 2: 1763-1767.

Ayalew W, Edward J, Rege O, Getahun E, Tibbo Mand Mamo Y. 2003, Delivering systematic information on indigenous animal genetic resources – the development and prospects of DAGRIS. In: Proc. Conference on Tropical and Subtropical Agricultural and Natural Resource Management (TROPENTAG) October 8-10, 2003, University of Gottingen, Deutschar.

Bhatia AK, Yadav DK and Vij PK. 2009. Flexible data processing system for field survey on domestic animal diversity. *Journal of Livestock Biodiversity* 1: 28-33.

Bhatia AK, Jain A, Sadana DK, Gokhale SB and Bhagat RL. 2010. Phenotypic identification of farm animal genetic resources using computer learning with scoring function. *Computers & Electronics in Agriculture* 73: 37-43.

Joshi BK and Vij PK. 2009. Conservation and utilization of animal genetic resources of India: present status and future perspective. In: Proc. National Symposium on Livestock Biodiversity Conservation & Utilisation: Lessons from Past and Future Perspectives, February 12-13, 2009, Karnal, India, pages 15-19.

Huang Y, Lan Y, Thamson SJ, Fang A, Hoffmann WC, Lacey RE. 2010. Development of soft computing and applications in agricultural and biological engineering. *Computers and Electronics in Agriculture* 71: 107-127.

NBAGR 2005. Project Co-ordinator's Report for 7th Scientist Meet of Network Project on Animal Genetic Resources. G. Singh, A. Jain, P. K. Singh (eds.), National Bureau of Animal Genetic Resources, Karnal, India.

NBAGR 2008. Project Co-ordinator's Report for 8th Scientist Meet of Network Project on Animal Genetic Resources. G. Singh, P. K. Singh, B. P. Mishra (eds.), National Bureau of Animal Genetic Resources, Karnal, India.