Design of intelligent decision support systems in agriculture

Petr Hanzlík, František Kožíšek, Josef Pavlíček

***Abstract* —** This article describes the role of decision support systems (DSS) in agriculture, with a special emphasis to the aspects of DSS architecture and design. Based on existing research studies, an overview of factors that are crucially important for acceptance of agricultural DSS is provided. Based on these premises, a client-server agent-based architecture of DSS is presented. The server part of the solution is discussed in details with respect to the knowledge model and knowledge engine construction and integration, and demonstrated on a case study of plant-leaf recognition agent. In relationship with identified lack of relevant training data for intelligent agents, a strategy for creating a feedback link is discussed, which enables the DSS to maintain its quality over the time, and continually improve its performance.

***Keywords*—** Decision Support Systems, Intelligent Agent, Knowledge engine, Agricultural DSS.

I. Introduction

Decision support systems (DSS) are nowadays widely used across many industries to support professional in the decision making process. A DSS can be used for supporting both structured and unstructured decision. However, it cannot fully replace the decision maker itself; DSS does not possess certain aspects of human decision making abilities – intuition, creativity, or imagination. However, it is good for extending decision maker's ability to process knowledge or to tackle complex problems. It can also save a lot of time which the decision maker would need to understand various aspects of the problem and process the information. It can help the decision maker to look out of the box and can make him aware of new approaches but on the other hand DSSs are not universal; they are usually narrowly focused to one topic, e.g. yield prediction, fertilizers. A DSS is also limited by the computer system it runs on, and by models and knowledge it is based on.

Marakas [1] defines five basic components of DSS:

- Data Management system used for storing and organizing the relevant data
- Model Management system used for performing the necessary analysis and modeling
- Knowledge engine used for problem recognition and problem solving
- User interface used for each of the previous components to make it accessible for its users
- Users who interact with the system are crucial from the system design point of view. DSS has to be adjusted to fit the knowledge, motivation and skills of the target user base.

Marakas [1] presents multiple classifications of DSS, too, focusing on several criteria. Based on its orientation, DSS can be classified either as data-centric or model-centric. While data-centric DSS is primarily focused on the data it processes, model-centric DSS is rather oriented on simulations and optimization tasks.

Another way to classify DSS corresponds to the way it is used by its end users – formal and ad-hoc systems. While formal systems are used on regular basis for recurring decision making, ad-hoc system are focused on helping to solve some immediate ad-hoc problems a decision maker faces. DSS can be classified into directed and non-directed systems, too. While directed DSS gives guidance to its user to structure and execute the decision making, the non-directed DSS stays on the other pole by not giving any guidance. Obviously DSSs are not black and white and one should rather talk about the degree of guidance DSS provides to the user.

Last division we mention here is on individual and group DSSs. Many decisions cannot be taken by individuals but rather should be taken by a group of specialists. The extent to which DSS support group decision is thus an important factor.

One of the key components of a DSS system is the knowledge upon which can be built the decisions. Most of the knowledge (inference) engines of DSS utilize some kind of structured knowledge, typically formalized as IF-THEN statements (Modus Ponens) [2]. Nevertheless, in the real world a lot of knowledge is stored in an unstructured form, e.g. documents written in a natural language. Therefore one of the important tasks in designing a DSS is to extract, organize and structure such knowledge in orders to make it available for fast and agile retrieval.

This paper was supported by the grant project of the Czech University of Live Sciences: Methodical framework for knowledge acquiring in agriculture, no. 20141051. This support is gratefully acknowledged.

Petr Hanzlík is with the Department of Information Engineering, Faculty of Economics and Management, Czech University of Life Sciences Prague, Kamýcká 961/129, 165 00 Prague 6-Suchdol, Czech republic, e-mail: hanzlikp@pef.czu.cz

František Kožíšek is with the Department of Information Engineering, Faculty of Economics and Management, Czech University of Life Sciences Prague, Kamýcká 961/129, 165 00 Prague 6-Suchdol, Czech republic, e-mail: kozisek@pef.czu.cz

Josef Pavlíček is with the Department of Information Engineering, Faculty of Economics and Management, Czech University of Life Sciences Prague, Kamýcká 961/129, 165 00 Prague 6-Suchdol, Czech republic, e-mail: pavlicek@pef.czu.cz

Ribino et al [3] researched the possibility to extract knowledge from written documents. He created a system based on two approaches – semantic space and ontology based model. Semantic space approach is based on analyzing distribution of words in text and their proximity, and identifying the topic of the text. Ontology approach in Ribino work on the other hand was not focused on the content of the text but on identification where the whole document comes from – in what project it was created, by whom and for what purpose.

In several cases, utilizing just the existing IF-THEN rules is insufficient to sufficiently explain the causes of variable variance. Human expert may be able to propose a satisfactory solution; however, often based on prior experience and intuition only. In situations when full understanding of strength of influence of exogenous variables on the final decision is not required, intelligent techniques, such as artificial neural networks [4], genetic algorithms and simulations [5], can be used to represent knowledge. In order for this technique to work reliably, several conditions must however be met, including a vast and representative data sample from the problem domain.

This article discusses construction of DSS with a special emphasis exerted towards aspects of user acceptance, and methods to design a knowledge model and engine combination (and the aspects of such systems in general). In relationship with intelligent aspects of a decision support system, the current lack of relevant data is discussed. Based on these findings architecture of DSS is proposed, and aspects of knowledge engines, data formalization, problem modeling, and acceptance among farmers discussed.

II. DSS IN AGRICULTURE

There has already been done a variety of research on decision support system in agriculture. Some of the authors are focused on one aspect of the agriculture, e.g. irrigation [6], rainwater [7] and water [8] management, dairy business [9] or crops managers [10] and analyze to what extent DSS can actually improve the way how farmers work

Other authors on the other hand are more focused on analyzing what general approach is the best for building DSS logic and models in order to get as accurate results as possible [11].

Even though we speak about DSS in agriculture in general, one should understand there is a big variety of different problems which can be handled by DSS – e.g. irrigation of fields, effective use of fertilizers, nutrition and feeding of cattle, products' price risk management, pesticide dosing etc. Each of these topics requires substantially different kind of expertise (domain knowledge) and a different approach and tools may be required to use this knowledge effectively. However, a success of DSS among farmers is not necessarily guaranteed by the architecture and specialization of tools used.

A. DSS acceptance aspects

Cabrera [9], who analyzed DSS used in dairy production business defined following five aspects of a successful DSS:

- Highly user-friendly
- Farm and user specific
- Grounded on the best scientific information available,
- Remaining relevant throughout time
- Providing fast, concrete, and simple answer to complex farmers' questions

As in DSS generally, the key is getting the right information to the right people in understandable form at the right time while not imposing additional workload on the farmers.

An important aspect for decision support system in agriculture is how to achieve acceptance among its users. There are multiple reasons which can prevent farmers to use the system.

As Chauhan et al. [6] points out, one of the reasons why farmers discontinued to use their system was that they considered it as too time consuming. The system for improved irrigation required its users to enter data about temperature every day. Although they saw added value in the system, the amount of work it required to keep the system relevant prevented them to use it. One of the options how to deal with this problem is to retrieve data from other public sources. Accordingly, Chauhan et al. [6] ware retrieving information about solar radiation and temperatures from meteorological stations.

Another important aspect for acceptance by end users can be the complexity of models with inability to access real time data [10]. This causes that even the farmer would like to use the system he cannot.

Other aspect is specialization of the system. As Chauhan et al. [6] claim one of the reasons their system did not get wider adoption was its specialization on peanuts. Farmers did not want to use multitude of systems for multiple products. Ease of use for farmers is thus critical.

Other aspect one should not forget is ICT technology farmers have available. Lack of broadband connection in some areas contributed significantly to limited adoption of DSS in case of Chauhan et al. [6] irrigation DSS. It is also important to note, that the workforce in agriculture generally have lover level of education than rest of the population, and the interest in undergraduate and graduate study programs in agriculture-related fields is gradually decreasing [13].

B. DSS architecture

There are different methods and tools used for creating DSS in agriculture. The key element is to choose a one that fits the nature of the task under review. Different problem generally require different set of knowledge and methods to solve them effectively and satisfyingly for the end users.

For purposes of irrigation DSS, GIS is a popular option. Triana et al [13] used a GIS based decision support system (GeoDSS) for automatic construction of flow networks which fits well for complex geo-databases with extensive spatial features. GIS was also used by Susilawati et al [7] for his model of rainwater management.

Yeping et al, [11] on the other hand suggests in his study usage of agent based decision system which is based on artificial intelligence research. He describes agents as a computer process which is able to adapt to the changing environment. He suggests to design DSS as a multi agent system, where each agent will have a different role (data management agent, model agent, interface agent, forecast agent etc.), but where the agents will collaborate with each other.

Keating et al. [14] describe in their research an approach used to develop their system APSIM. He recognizes the complexity an agriculture production brings, so APSIM was built as a general simulation framework, where individual modules model different agricultural production aspects. They are even developed by different groups and the framework interconnects them. Modules included in APSIM were for example crop growth simulation module, soil water balance module, soil organic matter, nitrogen or residues modules. These modules work as self contained “black boxes” and they communicate with each other through the central engine.

Cabrera [8] analyzed 30 different DSSs in dairy production. They were focused on different aspects of this activity – nutrition and feeding, reproduction efficiency, heifer management and cow replacement, production and productivity, price risk management, and environmental stewardship. Because of the variety of problems, different methodologies must have been used. For example, linear programming has been used in order to determine the cost effective diet composition. Partial budgeting and cost benefit was used in order to determine break-even for usage of feed additives. Mathematical simulation and projection was used to simulate nutrient requirements of each cow. Non-linear optimization was used in order to determine the minimum premium price to a defined target guarantee net income over feed cost according to future projected commodity prices and farm specific conditions. Main conclusion Cabrera [8] however did was that independently to the method used, the most important fact is to give to the farmer a tool, which will provide a solid, but still user-friendly instrument for their decision-making.

C. Problems with agricultural data

A very promising modular intelligent approach proposed by Yeping el al. [10] currently faces a significant obstacle – the lack of precise, representative, and up to date data and data sources in agriculture. This problem is not a trivial one, because in vast majority of problem domains, especially those approached on the add-hoc basis, a single agricultural enterprise is not capable of producing sufficient amount of data for training an intelligent agent, or even for drawing expert conclusions, that is.

Agricultural data are also mostly of temporal, spatial, or even spatio-temporal nature. To illustrate the problem, let us list a few examples here. A composition of pesticide agents evolves over time and pest can gain a resistance to certain compounds of these agents, therefore results from past may be representative no more. Certain agent may not be applied in proximity of water reservoirs or in protected natural zones to prevent containment. Impacts of a fertilizer may be influenced by soil-type (permeability), water content, or weather conditions at the moment of application, and in given amount of time before and after it. The list of such specific situation is immense.

Therefore, the only possible solution is sharing of data across the agricultural sector, which is obviously not natural for individual businesses, because of the possible loss of competitive advantage. However, the practice of sharing of best practices across the sector is gradually promoted by the public authorities in relationship with the long-term efforts to stabilize the agricultural sector and prevent ineffective agricultural treatments of the land, soil, and human resources.

III. Problem Solution

The previous sections of the article suggest that a multitude of conditions must be fulfilled in order to design a DSS with a potential to be accepted. The key problems to solve thus are:

- Problem modeling
- Knowledge engine design, structure and position
- Data acquiring strategy
- Data formalization and storage
- User-friendly design
- Knowledge sharing and distribution
- Continual improvement and actualization

To make the DSS as user-friendly as possible, it is important to eliminate the influence of significantly different user interfaces and user controls needed to accomplish decision-making tasks in individual sub-domains of agriculture. In optimal case, a single interface template could be used for every task. Another possible approach is to embed distributed knowledge engine into the structure of application farmers currently use for decision-making, and are thus familiar with. In order to provide a possible solution for both of these problems, we propose a client-server based DSS architecture, schematically presented in following figure:

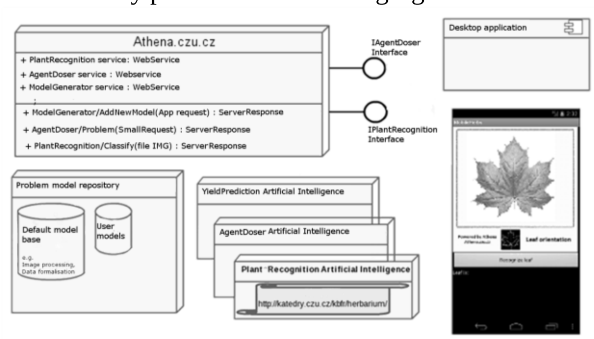

Fig. 1: Proposed architecture of agricultural DSS

It is not feasible to design a general-purpose DSS for agriculture with a single problem model and knowledge engine. However, an agricultural DSS can be realized using collection of problem models paired with corresponding knowledge engines (Fig. 2), referred to as intelligent knowledge modules (knowledge agents), in compliance with the proposal of Yeping [11]. In order to access individual knowledge modules, a web-service is constructed for each one of them, together with a class interface to submit required input data. Therefore a user interface of choice can be used, and an existing system can easily embed novel decision-making modules. The user-centric DSS is supported this way.

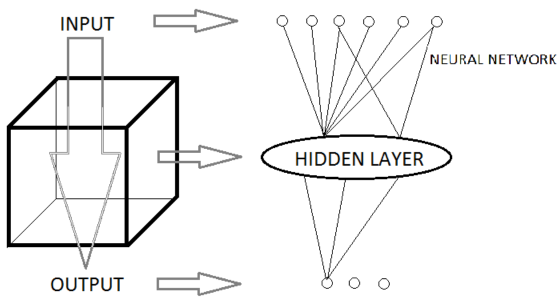

Fig. 2: Problem model to knowledge engine transition

A. *Case study – Leaf recognition knowledge module*

To illustrate the aspects of construction of knowledge module as a combination of problem model and knowledge engine, let us present an example describing the construction of an image-based leaf recognition knowledge module.

In order to construct a problem model, a significant amount of domain knowledge is required. In order to classify a leaf to a corresponding species from a basic input image (see Figure 3), several dimensions can be used, including the shape, color distribution, vein distribution, or combination of above mentioned methods.

Fig. 3: RGB input image – a maple leaf

In simplest case of problem module from the knowledge engineer perspective, a set of input directly classified images of plant species is fetched to a feed-forward neural network, with a size of hidden layer small enough to support generalization. An optimal knowledge engine architecture (in mentioned scenario corresponding to the architecture of neural network) may be acquired using a simulation scenario, which will test the performance of individual network models and architecture in order to select the most effective one for the problem, based on combination of error minimizing function, and convergence time.

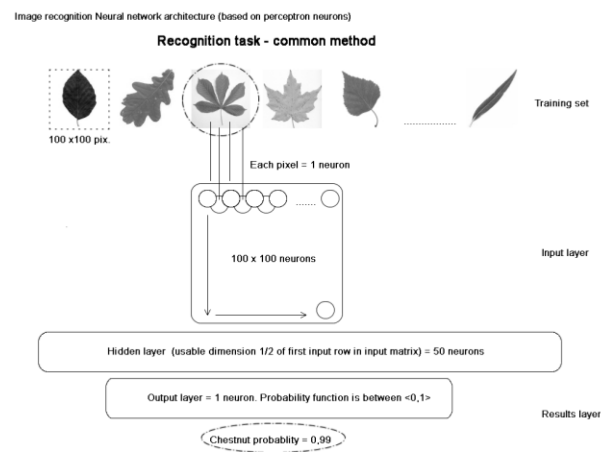

Fig. 4: Simple leaf recognition task

In such a case, the time required to train the artificial neural network will be very high, because of the dimensions of the input layer. The solution also will also be prone to the error, especially because of the image scaling and rotation. If such a solution were to work, a set of constrains must be set to ensure the proper image acquiring scenario for the model.

Therefore, a much more sophisticated problem models, based on a combination of several image characteristics can be constructed to overcome the above mentioned problems. To limit the input-layer network size, the image can be represented as a combination of color distribution histogram and shape features. Histogram can be calculated quite easily using a simple algorithm –see the MATLAB snippet in Figure 5 as a reference. The second characteristic of every leaf is its respective shape and veins distribution, which can be acquired using an edge detection algorithm. The shape characteristic can be further formalized using a distance of the edge from center of gravity in arbitrary number of circular cuts. This way, the model is not sensitive to image scaling, nor orientation.

```
%Specify the bin range[0 255]

bin=255;

%Find the histogram of the image.

Val=reshape(A,[],1);
Val=double(Val);
I=hist(Val,0:bin);


%Divide the result by number of pixels

Output=I/numel(A);

%Calculate the Cumlative sum
CSum=cumsum(Output);


%Perform the transformation S=T(R) where S and R in the range [ 0 1]
HIm=CSum(A+1);
```

Fig. 5: A problem model representation – MATLAB snippet, image histogram

The problem model thus significantly reduces the size of input vector (Figure 6), than in the original case, and the performance of the knowledge engine grows proportionally.

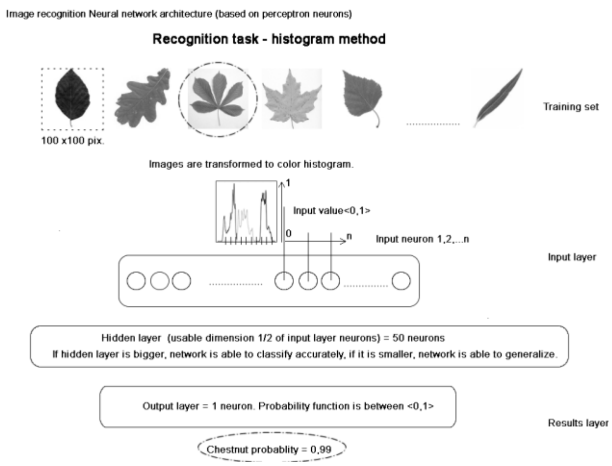

Fig. 6: Histogram based part of the leaf recognition method

In order to solve sophisticated tasks, the elementary components can be combined to create effective algorithmic preprocessing steps, which significantly improve the performance and accuracy of the decision-making task. This way, the histogram method can be combined with the shape-based method, and together provide rather reliable classification. From the end user point of view, the system yet only processes arbitrary image of a leaf, and provide a classification with a given level of certainty.

B. *Discussion*

Following the logic outlined above on a well established principle, we can conclude that in order to interpret agricultural data, a set of models must be designed and employed. These modules in such scenario are usually based on loosely formalized expert knowledge. After formalizing the knowledge in a form of intelligent system, the dimensionality of the input data can be reduced by intermediary problem models, or general techniques for the task, such as an identity networks, factor analysis or PCA. However, the more precise the knowledge model specification is, the more effective knowledge engine can be constructed.

Because of spatial, temporal and spatio-temporal character of data in agriculture, the knowledge models must also integrate several third party data sources. The GIS systems should be heavily relied upon to describe the terrain, based on the problem-specific intersections of available maps (geological maps, hydrological maps, terrain map, etc.), the past weather data, as well as future forecast information can be embedded into the input domain to optimize a decision-making procedure, based on available real-time data. The selection of relevant inputs depends primarily on the problem model; a formalized representation of the expert knowledge.

While creating a DSS at one given point of time is feasible, maintaining the continuous performance of the system is highly complicated task, because in many domains, the data which has been used as a training set are subject of temporal degeneration, and lose their relevance over the time. The continuous acquiring of fresh data is therefore vital. To accomplish this, user cooperation is required. Because of the high costs of employing experts, and because of the complex nature of many agricultural problems, small to medium sized agricultural enterprises often based their decision on limited (approximate) knowledge. The interest of these subjects to participate on developing an accessible yet robust solution is therefore logical. The advantages for big agricultural enterprises however require understanding of indirect implications of data driven knowledge engine design. Because the system heavily relies on past data and knowledge, adapting a new agent/technology/method requires a significant presence of such property in a training dataset. Therefore, producers of these materials need to focus on influential players first, to gain general acceptance. These business units can in fact play the role of a “knowledge maker”. The early adoption of new technologies into corporate practice can be reflected as a competitive advantage. The acquiring of real world data thus has been feasible for now.

IV. CONCLUSIONS

This article describes the specific aspects of decision support systems in agriculture with the help of previously conducted scientific studies within this domain. With the help of conclusions of these works, it presents a set of agriculture specific premises about the user base and nature of data in agriculture, which significantly influence the proposed framework for construction of agricultural DSS. The described modular client server architecture utilizes matching pairs of problem models and knowledge engine, to create sophisticated, specific problem focused knowledge models, which can be accessed using a web service. The proposed framework is demonstrated on the case study, which describe the simplified system for leaf recognition. Because of the data-driven nature of such DSS, it is however necessary to establish a continuous data acquiring strategy, that will allow retraining of the knowledge on the basis of up to date data. The knowledge acquiring phase of the DSS design requires much more discussion, especially from the point of view of creating a stable feedback link between users and the system, and from the point of view of exploiting the potential of opendata and public data sources.

REFERENCES

[1] G. M. Marakas, *Decision Support Systems in the 21st Century*. Prentice Hall, 2003, ISBN: 9780130922069.
[2] F. Hayes-Roth, D. Waterman, and D. Lenat, *Building Expert Systems*. Addison-Wesley, 1983, ISBN 0-201-10686-8.
[3] P. Ribino, et al., “A Knowledge Management and Decision Support Model for Enterprises“. *Advances in Decision Sciences*, 2011, vol. 2
[4] A. Fereidunian, et al., “Designing an intelligent decision support system for human-centered utility management automation part 1: structures, problem formulation, solution methodology”, *Proceedings of the 7th*

WSEAS international conference on Automatic control, modeling and simulation. 2005, pp. 505-510. ISBN:960-8457-12-2

[5] F. Stoica, L. F. Cacovean, “Using genetic algorithms and simulation as decision support in marketing strategies and long-term production planning”, *Proceedings of the 9th WSEAS international conference on Simulation, modelling and optimization.* 2009, pp. 435-439. ISBN 978-960-474-113-7

[6] Y. Chauhan, et al, “AQUAMAN: A Web-Based Decision Support System for Irrigation Scheduling in Peanuts”, *Irrigation Science*. 2013, vol. 31, no. 3, pp. 271-283. ISSN 03427188.

[7] C. L. Susilawati, S. TUNGGUL, “Model Development of Rainwater Management for Agriculture Decision Support System in Semi Arid Area”, *Civil Engineering Dimension*, vol. 13, no. 2, pp. 107-112. ISSN 14109530, 2011

[8] M. Cioca, L. Cioca, D. Cioranu, “Decision Support Systems used in Water Management”, *Proceedings of the WSEAS Int. Conf. on Waste Management, Water Pollution, Air Pollution, Indoor Climate*. 2007, pp. 208-211

[9] V. E. Cabrera, *DairyMGT: A Suite of Decision Support Systems in Dairy Farm Management*, 2012

[10] C. J. Fernandez, T.N. Trolinger, “Development of a Web-Based Decision Support System for Crop Managers: Structural Considerations and Implementation Case”, *Agronomy Journal*. 2007, ISSN 0002-1962.

[11] Z. H. U. Yeping et al., “Agriculture and Forestry Economy Decision Support System Based on Agent”, *New Zealand Journal of Agricultural Research*. 2007, vol. 50, no. 5, pp. 1339-1346. ISSN 00288233.

[12] O. Gojda, P. Hanzlik, T. Martinek, “The Career Perceptions of Future Graduates”, *Agrarian perspectives XXII: Development trends in agribusiness*, 2013, pp. 125-135

[13] E. Triana, J. W. Labadie, “GIS-Based Decision Support System for Improved Operations and Efficiency Conservation in Large-Scale Irrigation Systems”, *Journal of Irrigation & Drainage Engineering*. 2012, vol. 138, no. 10, s. 857-867. ISSN 07339437.

[14] B. A. Keating et al., “An Overview of APSIM, a Model Designed for Farming Systems Simulation”, *European Journal of Agronomy*. 2003, vol. 18, no. 3, pp. 267-288.