

UMA NOVA ABORDAGEM PARA DIAGNÓSTICO DE DOENÇAS VIA WEB¹

SILVIA MARIA FONSECA SILVEIRA MASSRUHA²

HELANO PÓVOAS DE LIMA³

RESUMO: Neste trabalho é apresentada e discutida a nova plataforma do módulo de diagnóstico de doenças (módulo produtor) do sistema Diagnose Virtual que foi desenvolvida para integrar ao portal WebAgritec. O WebAgritec é uma plataforma, concretizada na forma de um website com uma interface mais interativa para transferência das tecnologias da Embrapa para os extensionistas e produtores rurais. Esta plataforma, baseada nos conceitos de Web 2.0, visa o planejamento, previsão e monitoramento da produção agrícola e reúne informações sobre as culturas de arroz, feijão, milho, soja e trigo.

PALAVRAS-CHAVE: diagnose on-line, interface, interoperabilidade, doenças de plantas, web 2.0.

A NEW APPROACH FOR DIAGNOSIS OF DISEASES BY WEB

ABSTRACT: In this paper is presented and discussed a new architecture for the module of diagnosis of diseases (user module) of the Virtual Diagnosis System which was developed to be integrated in the portal WebAgritec. The WebAgritec is a platform, implemented in the form of a website with a more interactive interface to transfer technology from Embrapa to extension workers and farmers. This platform, based on the concepts of Web 2.0, aims the planning, forecasting and monitoring of agricultural production and gathers information about the cultivation of rice, beans, corn, soybeans and wheat.

KEYWORDS: diagnosis on-line, interface, interoperability, plant diseases, web 2.0.

1. INTRODUÇÃO

A Embrapa (Empresa Brasileira de Pesquisa Agropecuária), por meio de seus 45 centros de pesquisa, atua em parcerias na geração de tecnologias para os diversos segmentos sociais os quais trabalha, preservando e valorizando a biodiversidade e os recursos naturais visando uma agricultura sustentável. Todas as unidades da Embrapa, sejam elas da área de produção animal ou vegetal, procuram atuar nos pontos de estrangulamento de suas cadeias produtivas de modo a melhorar a qualidade e produtividade de seus produtos. Para tal, existem equipes interdisciplinares de pesquisadores em cada unidade, dentre eles, engenheiros, fitopatologistas, entomologistas e veterinários.

Apesar do grande volume de informações sobre a identificação de doenças e métodos de controle disponíveis na literatura especializada, muitas vezes produtores e técnicos agrícolas desconhecem os princípios básicos que devem nortear o uso racional dos agrotóxicos e se utilizam de um arsenal de produtos químicos para combater uma doença, com resultados duvidosos. Frequentemente, usados de maneira inadequada, estes produtos podem colocar em risco a saúde dos próprios aplicadores, além de causar danos ao ecossistema.

No acervo bibliográfico da Embrapa podem-se encontrar vários documentos sobre “Diagnóstico de Doenças e Métodos de Controle”. Entretanto, o conhecimento se apresenta disperso em documentos. Os especialistas, por sua vez, detêm o conhecimento e possuem atalhos que aprendem com a experiência e que os tornam ágeis em suas decisões. De modo a atender essa demanda foi proposto um projeto na Embrapa denominado “Infraestrutura virtual para apoio a tomada de decisão no diagnóstico e prognóstico de doenças de plantas”, conhecido como Diagnose Virtual (Massruhá et al, 2007). Este projeto visou desenvolver uma infraestrutura única na área de sanidade para

¹ Financiada pela carteira do Macroprograma 2 da Embrapa e PAC/Embrapa

² Pesquisador da Embrapa Informática Agropecuária, CP 6041, 13083-886, Campinas, SP (silvia@cnptia.embrapa.br)

³ Analista da Embrapa Informática Agropecuária, CP 6041, 13083-886, Campinas, SP (helano@cnptia.embrapa.br)

diagnóstico e prognóstico de doenças via internet. O foco de atuação deste projeto foi desenvolver um sistema para auxiliar fitopatologistas e agricultores no diagnóstico e prognóstico de doenças causadas por fatores bióticos (fungos, bactérias, vírus e nematoides) e abióticos (fatores ambientais). As culturas que foram tratadas no escopo deste projeto foram divididas em 3 grupos: gramíneas (milho, trigo), leguminosas (feijão, soja) e hortaliças (tomate e pimentão). Posteriormente também foram integradas as culturas de arroz, coco, uva e pimenta-do-reino.

Na área de diagnóstico foram utilizadas técnicas de inteligência computacional tais como lógica abdutiva e nebulosa (Massruhá et al., 2005). Métodos mais adequados para aquisição de sintomas utilizando técnicas de reconhecimento de padrão também foram estudados. No escopo de previsão de doenças foi investigada uma abordagem integrada de técnicas de raciocínio incompleto e impreciso (lógica nebulosa) e técnicas de aprendizado de máquina (árvores de decisão, redes neurais). Ferramentas de geoprocessamento também foram integradas para auxiliar na geração de relatórios gerenciais na área de previsão de doenças. Este projeto também visou disponibilizar estes serviços de diagnóstico e alerta de doenças em celulares, palmtops e pagers, utilizando técnicas de computação móvel e possibilitando os agricultores e técnicos acessá-los mesmo quando não for possível utilizar um computador na internet. Esta infraestrutura se deu na forma de uma aplicação web e foi desenvolvida em Java utilizando tecnologias J2EE (Java 2 Platform, Enterprise Edition). Embora esta tecnologia tenha se mostrado adequada, ela apresentava problema de desempenho. Diante disto, foi proposto um novo módulo de diagnóstico baseado na tecnologia Web 2.0 de modo a integrar ao portal Webagritec.

O WebAgritec é uma plataforma que visa o planejamento, previsão e monitoramento da produção agrícola e reúne informações sobre as culturas de arroz, feijão, milho, soja e trigo. O sistema apresenta ao usuário, diferentes módulos: época de plantio com menores riscos climáticos associados; cultivar mais apropriada para os propósitos do usuário; indicação de calagem e adubação para cada cultura; previsões e tendências das condições climáticas antes e durante a safra; diagnóstico de doenças e distúrbios nutricionais que possam aparecer no decorrer da safra. O foco deste trabalho é apresentar a nova plataforma do módulo de diagnóstico de doenças.

A validação deste trabalho foi feita através do desenvolvimento de aplicações sob esta infraestrutura. Foram previstas aplicações para diagnóstico de doenças das 5 culturas disponíveis no portal Webagritec. Cabe ressaltar ainda que, neste primeiro momento, foi contemplado apenas o manejo de doenças conforme apresentado na seção 2. Entretanto, a tecnologia desenvolvida servirá de base para futuras ações nas áreas de controle de plantas daninhas e o manejo integrado de pragas causadas por insetos e ácaros. Na seção 3, são apresentados os materiais e métodos utilizados no escopo deste trabalho. Na seção 4, estão descritos os principais resultados obtidos neste trabalho. Finalmente, na seções 5 e 6, são apresentados as conclusões bem como os trabalhos futuros e as referências bibliográficas.

2. OBJETIVO

Este trabalho tem por finalidade apresentar a nova arquitetura proposta para disponibilizar o módulo de diagnóstico de doenças para o portal Webagritec, que é uma plataforma que visa o planejamento, previsão e monitoramento da produção agrícola. Neste trabalho também são discutidas as principais diferenças entre a nova arquitetura e a anteriormente utilizada no sistema de Diagnose Virtual. Funcionalidades adicionais foram incorporadas nesta nova versão do sistema para atender a requisitos de desempenho e interface do projeto. A nova interface, mais interativa, é apresentada detalhadamente visando mostrar como os requisitos de interatividade contribuem para melhorar a qualidade da informação apresentada ao usuário, resultando em um diagnóstico mais eficiente.

3. MATERIAL E MÉTODOS

O Sistema Diagnose Virtual dividi-se em dois módulos principais, o módulo do especialista e o módulo do produtor. O primeiro tem a finalidade de cadastrar as bases de conhecimento sobre os

sintomas e doenças das culturas. Atualmente este módulo possui registros detalhados sobre as culturas de soja, milho, feijão, arroz, trigo, tomate, pimentão, videira, pimenta do reino, coqueiro e cana-de-açúcar, sendo esta última em fase de construção. O segundo módulo efetua o diagnóstico seguindo uma ordem específica. Nele o usuário começa indicando as características de sua propriedade, inclusive sua localização geográfica, passando pelos dados da cultura, data da observação dos sintomas e partes afetadas da planta. Posteriormente são apresentadas imagens para que ele identifique uma que se assemelhe ao problema e então são feitas perguntas associadas à imagens de sintomas conforme as possibilidades inferidas. Conforme o usuário responde as perguntas, o sistema apresenta as possíveis doenças.

Motivado pela necessidade da integração do sistema diagnose virtual ao portal Webagritec, e devido a diferença entre as tecnologias adotadas, foi necessário adequar o módulo do produtor às tecnologias e aos requisitos de desempenho e interatividade adotados pelo Webagritec. Um primeiro protótipo do Webagritec foi desenvolvido em Openlaszlo (<http://www.openlaszlo.org>) para consolidação dos conceitos do novo portal e definição de tecnologia utilizada na camada de visão. Além do Openlaszlo, outras tecnologias na Web 2.0 como JAVA FX e Adobe Flex foram estudadas e, optou-se por desenvolver o novo portal do Webagritec em Adobe Flex® (ADOBE, 2010).

A versão anterior do sistema diagnose virtual, possui uma arquitetura baseada no Java EE (JSP / Servlets) seguindo o padrão MVC (*Model-View-Controller*) que visa separar as tarefas de acesso aos dados e lógica do negócio, da apresentação e interação com o usuário, introduzindo um componente entre os dois: o *Controller*. Alguns *frameworks open source* bem segmentados no mercado foram utilizados na construção do sistema anterior, entretanto, apesar de se tratar de ferramentas bastante completas e robustas, a adoção destas ferramentas gera uma sobrecarga no código e uma grande dificuldade de customização, pois são desenvolvidas focando sistemas de informação bem mais complexos, onde é justificável o uso de todas as funcionalidades e facilidades que elas oferecem.

Particularmente, o framework de persistência objeto relacional adotado, mostrou-se inadequado perante a característica peculiar do sistema de manipular imensas hierarquias de objetos em memória durante a inferência de possibilidades, chegando a realizar durante testes a incrível quantidade de 20 mil consultas ao banco de dados.

Para implementação do sistema, todo o código original na linguagem Prolog foi traduzido para o Java, o que mostrou-se também uma alternativa não viável do ponto de vista de performance, já que Prolog é orientado à um paradigma declarativo e de lógica matemática, baseando-se principalmente em recursão, enquanto Java segue o paradigma estruturado e orientado à iteração.

Além dos problemas de desempenho citados, a maneira como o sistema havia sido codificado, seguindo uma ordem fixa de passos e das perguntas feitas ao usuário, não era compatível com o paradigma do Webagritec, que tenta levar, o mais próximo possível, o comportamento de uma aplicação desktop, onde o usuário clica aonde quer, no momento que deseja.

Foi então proposta uma nova plataforma para o sistema afim de sanar ambos os problemas. A arquitetura continuou baseando-se no modelo MVC e na arquitetura Java EE na parte servidora, entretanto no lugar do uso de frameworks, os componentes controlador e de persistência de objetos foram escritos especificamente para o sistema e para a camada de visão foi adotada a plataforma Adobe Flex®.

Para o componente de inferência, optou-se por construir um gerador de código Prolog em Java, que dinamicamente gera a base de conhecimento para esta linguagem e executa a inferência seguindo o algoritmo original em um interpretador Prolog e devolve o resultado para o Java. Vários interpretadores foram testados para incorporação, dentre eles o GNU Prolog (GNU, 2010), SWI Prolog (SWI, 2010) e o Amzi! Prolog (AMZI, 2010) mas o que obteve de longe o melhor resultado foi o YAP Prolog (YAP, 2010) desenvolvido na Universidade do Porto especificamente focando alta performance. Visualizando a Tabela 1, podemos ter uma visão comparativa das duas arquiteturas:

Componente	Sistema Original	Nova Versão
Camada de visão	JSP, Sitemesh, Displaytag, Struts Validator	Adobe Flex®
Camada de controle	Struts	Componente próprio otimizado
Camada de Persistência	Spring, Hibernate	Componente próprio otimizado
Banco de Dados	PostgreSQL	PostgreSQL
Motor de inferência	Escrito em Java	Gerador de código Prolog dinâmico + interpretador Prolog YAP

Tabela 1 – Comparativo de arquiteturas

4. RESULTADOS E DISCUSSÃO

Na Tabela 2 são apresentados os resultados da comparação do tempo de execução das camadas de inferência e persistência entre a versão original e a versão atual. O tempo médio em milissegundos de 10 execuções em um CPU AMD64 Quad core com 8GB de RAM, usando o profiler/debugger do Netbeans no SO Linux Ubuntu 10.10 e a cultura do Milho para o primeiro ciclo de diagnóstico:

Camada	Versão Anterior	Versão Nova
Inferência	2210ms	132ms
Persistência	13324s	334ms

Tabela 2 – Comparativo do tempo de execução

As demais camadas é inviável comparar o tempo pois as arquiteturas e a dinâmica do diagnóstico são diferentes (JSP vs Flex que é assíncrono), mas elas não tem influência alguma no tempo de execução.

Devido a reestruturação do módulo do produtor do sistema Diagnose Virtual descrita anteriormente, a nova interface, que pode ser vista nas Figura 1 e 2, teve sua performance, apresentação e eficiência bastante melhorada.

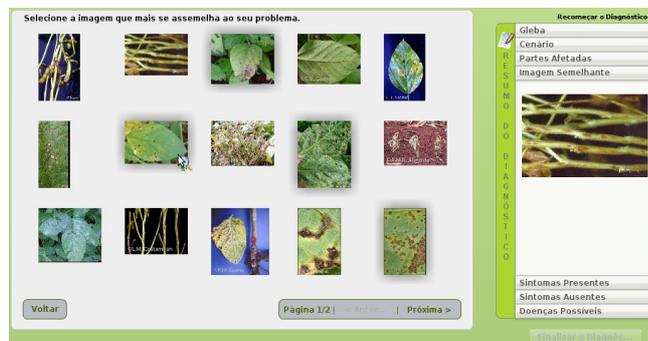


Figura 1 – Nova interface do sistema Diagnose Virtual



Figura 2 – Componente interativo dos sintomas

Nela o usuário pode retornar ao passo anterior do diagnóstico a qualquer hora, bem como escolher quais perguntas quer responder dentre as oferecidas e fazê-lo em qualquer ordem através de um componente gráfico de carrossel onde clicando-se na imagem a pergunta correspondente é mostrada. A quantidade e qualidade da informação oferecida ao usuário foi outro grande ponto de melhoria. Foram adicionados componentes para descrever e mostrar graficamente as fases de desenvolvimento e as partes das plantas, bem como, a cada passo, um relatório dos dados do diagnóstico é mostrado lateralmente, oferecendo fácil acesso às informações.

Foi introduzido também, no momento em que o sistema encontra alguma doença possível, um componente em formato de livreto informativo (Figura 3), com todos os detalhes acerca daquela doença, como descrição e imagem dos sintomas, condições favoráveis, exames e medidas de

controle possíveis. Este formato é mais amigável ao usuário, pois além das características de interatividade, assemelha-se aos livretos em papel editados pela Embrapa para tal finalidade.

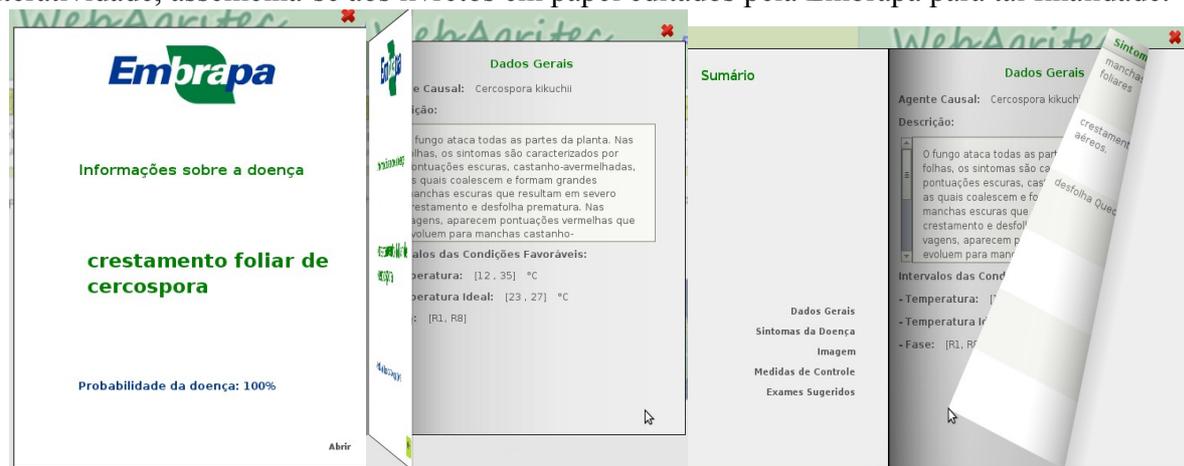


Figura 3 – Livreto interativo sobre a doença

5. CONCLUSÕES

O novo módulo do produtor do sistema Diagnose Virtual, além de ter gerado uma interface mais amigável para o usuário, proporcionou também um drástico ganho de desempenho devido à reimplementação do motor de inferência e a substituição dos frameworks da versão anterior por componentes próprios otimizados. Segundo testes, na versão anterior, para cada interação durante o diagnóstico, era necessário esperar pela resposta entre 45 e 90 segundos, conforme o tamanho do universo de possibilidades. Com a nova implementação, este tempo de resposta baixou para no máximo 2 segundos, mostrando-se plenamente adequado às características dinâmicas da nova interface. É válido lembrar ainda que as duas interfaces acessam o mesmo banco de dados independentemente garantindo a interoperabilidade do sistema. Como trabalhos futuros, pretende-se disponibilizar este módulo de diagnóstico de doenças em celulares, palmtops e pagers, utilizando técnicas de computação móvel e possibilitando os agricultores e técnicos acessá-los remotamente independente de um computador na internet.

6.REFERÊNCIAS

- ADOBE – Adobe Flex®. Framework to build internet applications. <>. Acesso <http://www.adobe.com/products/flex/> em jan. 2010
- AMZI INC. - Comercial prolog interpreter from Amzi! Inc. 2010. Disponível em: <<http://www.amzi.com/>>. Acesso em jul. 2010.
- GNU Prolog - Open Source prolog interpreter from GNU. 2010. Disponível em: <<http://www.gprolog.org/>>. Acesso em ago. 2010.
- MASSRUHÁ, S. M. F. S. ; SANDRI, S. ; WAINER, J. ; MORANDI, M. **An integrated framework for clinical problem solving in agriculture.** In: Efita /WCCA 2005 Joint Congress on IT in Agriculture, 2005, Vila Real. Efita /WCCA 2005 Joint Congress on IT in Agriculture, 2005a. p. 1400-1407.
- MASSRUHÁ, S. M. F. S. ; Dutra, J. P. ; CRUZ, S. A. B. ; SANDRI, S. ; WAINER, J. ; MORANDI, M. . An objected oriented framework fo virtual diagnosis. **In: 6th Biennial Conference of the European Federation of IT in Agriculture - Efita 2007, Glasgow.** EFITA/WCCA 2007 6th Biennial Conference of the European Federation of IT in Agriculture. Glasgow - Escócia : Glasgow Caledonian University, 2007.
- SWI Prolog - Open Source prolog interpreter licensed under LGPL. 2010. Disponível em: <<http://www.swi-prolog.org/>>. Acesso em jul. 2010.
- YAP Prolog - High-performance Prolog compiler from LIACC/Universidade do Porto. 2010. Disponível em: <<http://www.dcc.fc.up.pt/~vsc/Yap/>>. Acesso em ago. 2010.