

CADASTRAMENTO DE COLETAS DE FORMIGAS NO CAMPO.

Camila Kamarad Zocal Garcia, Carlos Norberto Fischer, Delano Medeiros Beder, Maurício Bacci Júnior. – Ciência da Computação – Bacharelado em Ciências da Computação – Departamento de Estatística, Matemática Aplicada e Computação – Instituto de Geociências e Ciências Exatas – Campus de Rio Claro.

Um Laboratório de Evolução Molecular (LEM) desenvolve pesquisas voltadas para o estudo de sistemática e evolução molecular com o objetivo de buscar relações entre espécies de seres vivos (Swofford, 2000). Os LEMs executam uma rotina de trabalho que envolve, basicamente, as seguintes etapas (Fischer et al, 2002): (1) coleta de organismos de interesse, (2) extração do material genético, (3) geração de cópias do material genético, (4) purificação do material, (5) clonagem para produção de cópias de cada sequência, (6) purificação do processo de clonagem, (7) sequenciamento do material genético e alinhamento das sequências e (8) estudos filogenéticos (geração de árvores filogenéticas). Parte dessas atividades, executadas no fluxo de trabalho de um LEM, consistem em alimentar bancos de dados e programas de Bioinformática com dados coletados no campo ou gerados nos experimentos.

As etapas descritas acima são realizadas sequencialmente pelos pesquisadores de um LEM, gerando um grande volume de dados em diferentes formatos. Assim, torna-se importante o desenvolvimento de uma forma de automatizar, quando possível, tais etapas, tanto para torná-las mais fáceis e rápidas quanto para evitar que algum dado importante deixe de ser obtido e armazenado no banco de dados (Fischer et al, 2002).

O Grupo de Bioinformática de Rio Claro (GBIRC) desenvolveu um sistema computacional, denominado SiGLEM – Sistema para Gerenciamento de Laboratório de Evolução Molecular, voltado para o gerenciamento e a automação do fluxo de trabalho em LEMs (Costa & Fischer, 2005). Atualmente, o SiGLEM está sendo utilizado pelo LEM do Centro de Estudo de Insetos Sociais (LEM/CEIS) do Instituto de Biociências de Rio Claro dentro do projeto denominado Formigas *Attini*. Para este projeto, o SiGLEM apresenta vários formulários *web* direcionados para o cadastramento (a inserção) numa base de dados relacional de um conjunto enorme de informações sobre formigas, que vão desde informações sobre a coleta deste tipo de organismo (como local da coleta, data, coletor, temperatura e vegetação associada) até dados moleculares obtidos no LEM/CEIS (como dados de extração, amplificação, clonagem e sequenciamento do material genético). Esta base de dados é alimentada com informações do tipo texto e com referências a diretórios que contêm arquivos (que também são submetidos através dos formulários) de vários tipos, como texto, tabelas, gráficos e fotos de géis. Vários outros formulários *web* possibilitam que tarefas de alteração e consulta das informações armazenadas na base de dados também possam ser realizadas remotamente. Devido ao fato de ser um sistema voltado para a *web*, o SiGLEM permite que os dados armazenados sejam compartilhados por grupos de pesquisa distintos.

No entanto, o SiGLEM não pode ser utilizado em locais onde haja dificuldade de acesso à Internet, por exemplo, no campo, usualmente o local de coleta de formigas, o que acaba levando os pesquisadores que fazem as coletas a registrarem os dados relacionados em planilhas eletrônicas, criando, assim, a possibilidade da não obtenção e/ou registro de alguma informação importante no momento da coleta.

Para resolver estes problemas, está em fase final de testes um outro sistema, desenvolvido pelos autores chamado SiGLEM *off-line*, que deverá ser instalado nos computadores portáteis dos pesquisadores que vão para o campo realizar coletas de formigas. O SiGLEM *off-line* é direcionado exclusivamente para o cadastramento de dados de coleta (ele não é voltado para alterar ou consultar a base de dados remota do projeto Formigas). O sistema deverá fazer com que todas as informações necessárias referentes às coletas sejam obtidas no momento da realização das mesmas.

O SiGLEM *off-line* foi implementado seguindo as mesmas especificações da versão *web*, mantendo, com isso, a compatibilidade (quanto a tipos e formatos) na obtenção dos dados a serem armazenados na base e facilitando a familiarização dos usuários com dois ambientes de trabalho que apresentam interfaces com formulários bastante similares.

O SiGLEM *off-line* funciona da seguinte forma. Vários formulários, desenvolvidos na linguagem Java, são apresentados ao pesquisador que está fazendo a coleta, para o

cadastramento de todas as informações necessárias. A Figura 1 mostra um formulário do SiGLEM que permite o cadastramento das formigas.

The screenshot shows a web browser window with the title bar. The page header has a logo with the word "FORMIGA" and a sun-like icon, and text "LEM - CEIS" and "Unesp Rio Claro". Below the header is a navigation bar with three tabs: "Cadastro", "Consulta", and "Alteração". The "Cadastro" tab is active, showing the "Cadastramento de Ninho" form. The form has a title "Cadastramento de Ninho" and a subtitle "Forneça as informações referentes à classificação do ninho:". It contains several dropdown menus: "Ordem" (Hymenoptera), "Família" (Formicidae), "Subfamília" (Myrmicinae), "Gênero" (Mycetarotes), and "Subgênero" (Não Especificado). There are also buttons for "Cadastrar novo gênero" and "Cadastrar nova espécie". Below these are fields for "Espécie" and "Subespécie" with a "Subespécie" dropdown menu (Item 1) and buttons for "Cadastrar nova subespécie". At the bottom are "Submeter" and "Interromper Cadastro" buttons.

Figura1 - Formulário para cadastro das formigas.

Basicamente, são cadastrados dois tipos de informação referente à coleta realizada: (1) dados do ninho e (2) dados da coleta. Para o cadastro do ninho, é necessário que o usuário faça a inserção, dentre outras, de informações referentes à localização geográfica do ninho, à vegetação associada ao local e à identificação taxonômica (ordem, gênero, espécie, etc.) das formigas do ninho (como pode ser visto na Figura 1). Para o cadastramento de uma coleta, informações como, por exemplo, tamanho do ninho na ocasião da coleta, temperatura, pesquisadores que realizaram a coleta, data, ferramenta de coleta e número de castas referentes às formigas coletadas são importantes. Essas informações são solicitadas em formulários criados de acordo com a natureza de cada uma delas. A Figura 2 mostra um formulário do SiGLEM *off-line* que permite cadastrar dados específicos de uma coleta de formigas.

Ao final do cadastramento de cada coleta, o sistema gera automaticamente um arquivo XML com os dados cadastrados. No arquivo XML gerado, estas informações são inseridas entre marcadores de texto específicos que permitem a fácil e rápida extração dos dados de interesse. Este formato mostrou-se adequado para o sistema devido à possibilidade de utilização de marcadores que indicam exatamente o tipo da informação armazenada. Este formato também é adequado para troca de informações via Internet, o que acontece na fase de submissão de cada arquivo XML ao servidor *web* do GBIRC, para a inclusão das informações da coleta na base de dados do projeto. A Figura 3 mostra um exemplo de trecho de um arquivo XML gerado pelo SiGLEM *off-line* onde podem ser observados os marcadores específicos para cada tipo de dado cadastrado.

The screenshot shows a web browser window with the title bar. The page header has a logo with leaves and the word "FORMIGA" in large letters, followed by a sun icon. To the right, it says "LEM - CEIS" and "Unesp Rio Claro". Below the header is a navigation bar with three tabs: "Cadastro" (highlighted in orange), "Consulta", and "Alteração". The main content area is titled "Cadastramento de Coleta". It contains several form fields and instructions. On the left, there are three dropdown menus for "Pesquisador 1" (selected: Alexander), "Pesquisador 2", and "Pesquisador 3". Below them is a text field for "Vegetação que as formigas estão cortando". On the right, there is a text field for "Ferramenta utilizada para a coleta", a dropdown for "Nº de castas" (selected: 12), a button "Castas e nº de formigas", a text field for "Observação referente às castas de formigas coletadas", and a text field for "Caixa onde o lote foi armazenado". At the bottom right, there is a date selection section "Data da Coleta" with "Dia" (4), "Mes" (Janeiro), and "Ano" (2006). At the bottom left, there are two buttons: "Submeter" and "Interromper cadastro".

Figura 2 - Formulário para cadastro da coleta.

```
<?xml version="1.0" encoding="UTF-8" ?>
<Ninho>
  <CodigoNinho>111</CodigoNinho>
  <Pais>Brasil</Pais>
  <Estado>São Paulo</Estado>
  <Cidade>Rio Claro</Cidade>
  <TipoVegetacao>Cultivodecafé</TipoVegetacao>
  <SubtipodeVegetacao>Não especificada</SubtipodeVegetacao>
  <Ordem>Hymenoptera</Ordem>
  <Genero>Mycetarotes</Genero>
  <Familia>Formicidae</Familia>
  <Subfamilia>Myrmicinae</Subfamilia>
  <Profundidade>1</Profundidade>
  <temperaturaNinho>1</temperaturaNinho>
  <Numerodeentradas>1</Numerodeentradas>
  <TemperaturaAmbiente>10°</TemperaturaAmbiente>
  <Numerodecamaras>1</Numerodecamaras>
  <Tamanhodoninho>2</Tamanhodoninho>
</Ninho>
```

Figura 3 - Arquivo XML gerado pelo SiGLEM *off-line*.

Os arquivos XML ficam armazenados nos computadores portáteis dos pesquisadores que realizam as coletas, que, quando tiverem acesso à Internet, poderão submeter tais arquivos ao servidor

web do GBIRC. Estão sendo desenvolvidas, pelos autores, ferramentas computacionais voltadas para a recepção e tratamento dos arquivos XML que chegam ao servidor *web*, para a extração dos dados e inclusão dos mesmos na base do projeto Formigas (algumas dessas ferramentas já se encontram em fase de testes, mostrando-se adequadas e de funcionamento correto).

Testes iniciais realizados pelos pesquisadores do LEM/CEIS têm demonstrado o adequado funcionamento do SiGLEM *off-line*. Este sistema tem se apresentado como uma ferramenta auxiliar bastante importante para uso dos pesquisadores no momento da coleta, pois espera-se, com este sistema, que sejam obtidas todas as informações necessárias sobre uma coleta ainda no campo, no momento da realização da mesma, evitando a possível necessidade de retorno ao local de uma coleta, às vezes em regiões bastante distantes do CEIS e/ou de difícil acesso.

Após a finalização do SiGLEM *off-line*, cópias do mesmo serão disponibilizadas a todos os pesquisadores que trabalham no projeto Formigas *Attini* para uso durante suas viagens de coleta de formigas.

Referências Bibliográficas:

- Costa, E.P; Fischer, C.N., 2005 - Relatório Técnico de Bolsa de Iniciação Científica da Fapesp. Fevereiro.
- Fischer, C.N.; Guilherme, I.R.; Bacci Jr., M., 2002 - “Uma Arquitetura de Bioinformática para Automação de Laboratórios de Evolução Molecular”. International Information Technology Symposium, Florianópolis, SC, Brasil.
- Swofford, D.L., 2000 - “Phylogenetic Analysis Using Parsimony, version 4.0b4a”. Illinois Natural History Survey, Champaign.

Bolsa: Fapesp