

SISTEMA DE RECONHECIMENTO IMUNOLÓGICO ARTIFICIAL PARA IDENTIFICAÇÃO DE OPERAÇÕES DE PERFURAÇÃO EM POÇOS PETROLÍFEROS

ADRIANE B. DE S. SERAPIÃO¹, JOSÉ RICARDO P. MENDES², KAZUO MIURA³

¹DEMAC/IGCE/UNESP
C.P. 178, Rio Claro, São Paulo, Brasil
adriane@rc.unesp.br

²UNICAMP/FEM/DEP
C.P. 6052, Campinas, São Paulo, Brasil
jricardo@dep.fem.unicamp.br

³PETROBRAS, E&P-SSE/UN-BC/ST/EP
Macaé (RJ), Brazil
kazuo.miura@petrobras.com.br

Abstract— *This work presents a artificial immune system used to automatically classify the petroleum well drilling operation stages through the analysis of some mud-logging parameters. A classification elaborated by a Petroleum Engineering expert was compared with the immune system classification in order to evaluate its performance.*

Keywords— artificial immune systems, classification, pattern recognition, drilling, mud-logging data, petroleum engineering.

Resumo— Este trabalho apresenta um sistema imunológico artificial para classificar automaticamente o estágio da operação de perfuração de poços de petróleo através da avaliação de alguns parâmetros de registros de *mud-logging*. Uma classificação realizada por um especialista da área de Engenharia de Petróleo foi comparada com a classificação resultante do sistema imunológico para validar sua eficiência.

Palavras-chave— sistemas imunológicos artificiais, classificação, reconhecimento de padrões, perfuração, *mud-logging*, engenharia de petróleo.

1 Introdução

A exploração de petróleo no mar avançou substancialmente nos últimos anos. O desenvolvimento e aprimoramento das técnicas de perfuração têm proporcionado a exploração de hidrocarbonetos em horizontes cada vez mais profundos, envolvendo altos custos e exigindo alta qualificação da equipe executora. Nos últimos anos, os avanços tecnológicos dos procedimentos de perfuração contribuíram para diminuir custos e para expandir as áreas de exploração. Parte do progresso tecnológico na área de Engenharia de Petróleo foi motivada pela evolução das técnicas de instrumentação. O desenvolvimento de sensores e sistemas de processamento de dados permitiu medir e registrar informações que abrangem desde parâmetros operacionais da sonda até características geológicas da formação perfurada. Com o desenvolvimento de tecnologias de transmissão de dados, essas informações, cujo acesso inicialmente era restrito ao ambiente da sonda, podem ser acessadas *on-line* de qualquer parte do mundo (Tavares *et al.*, 2005).

Como resultado do aumento do nível de instrumentação, existe uma grande quantidade de dados sendo medidos e registrados. Porém, as técnicas de análise de dados de perfuração não se consolidaram o suficiente para acompanhar o aumento na disponibi-

lidade de dados. As ferramentas desenvolveram-se bastante, mas no contexto da perfuração de poços, a análise de dados ainda ocupa uma posição de pouco destaque.

Este trabalho propõe um sistema de classificação automático que usa alguns parâmetros medidos e registrados durante operações de perfuração para gerar um registro preciso dos estágios da operação de perfuração de poços de petróleo. Duas aplicações são possíveis para este sistema de classificação. A primeira aplicação refere-se à análise de desempenho. Neste sentido, esta ferramenta pode ser utilizada para se fazer uma análise *a posteriori* do tempo gasto para perfurar cada poço de um campo de petróleo e investigar o tempo consumido por cada estágio durante a perfuração completa. A segunda aplicação está relacionada com a produção de um registro *on-line* dos estágios executados durante a perfuração.

O presente artigo apresenta um classificador imunológico artificial limitado a recursos, inspirado em mecanismos exibidos em sistemas imunológicos biológicos e artificiais, tais como seleção clonal, maturação de afinidade, retenção de células de memória, células B, competição por recursos e discriminação própria/não-própria (de Castro & Timmis, 2002). Este paradigma é a base de desenvolvimento do classificador automático para a identificação das opera-

ções de perfuração em poços de petróleo. Para realizar esta tarefa foi escolhido o algoritmo AIRS.

O Sistema de Reconhecimento Imunológico Artificial (AIRS) é um algoritmo de reconhecimento de padrões proposto por Andrew Watkins (2001), baseado em uma abordagem inspirada em princípios da imunologia. Os algoritmos que usam sistemas imunológicos artificiais são inspirados em populações de agentes treinados para reconhecer certos aspectos de interesse em um dado domínio. Neste trabalho utilizou-se uma versão melhorada do algoritmo AIRS original, denominada AIRS2 (Watkins & Timmis, 2002) com uma extensão para uma abordagem de processamento paralelo (Watkins & Timmis, 2004), que será referida como *AIRS2 Paralelo*.

Dados reais de um poço de petróleo capturados pela ferramenta de coleta de dados conhecida por *mud-logging* são usados para avaliar o desempenho do classificador AIRS2 Paralelo. Os resultados obtidos são comparados com a classificação prévia de tais dados por um especialista (Tavares *et al.*, 2005).

2 Sistema de Mud-logging

Durante a operação de perfuração de um poço de petróleo muitos parâmetros hidráulicos e mecânicos são medidos e monitorados para a execução da perfuração de maneira segura e otimizada. Existem diversos sistemas que atuam juntos em uma sonda de perfuração para realizar esta tarefa. Um destes sistemas, chamado de sistema *mud-logging*, é responsável pela medida de um conjunto de parâmetros mecânicos e geológicos. Com o desenvolvimento das técnicas de instrumentação, o número de parâmetros medidos aumentou e o uso do sistema de *mud-logging* tornou-se uma prática comum na indústria de petróleo.

Os sistemas de *mud-logging* dispõem de um grande número de sensores distribuídos nos sistemas operativos da sonda de perfuração. Uma característica desse sistema de informações é que nenhum sensor fica instalado dentro do poço que está sendo perfurado, ou seja, todas as medidas são realizadas na própria sonda. Os dados coletados pelos sensores são enviados para um sistema de computação central, onde os dados são processados e visualizados em tempo real em monitores instalados na cabine de *mud-logging* e na sala do supervisor da operação.

A monitoração da evolução dos parâmetros medidos é realizada, durante toda a perfuração, por um operador que deve disparar um alerta em caso de observação de alguma anormalidade. Se um parâmetro apresenta um comportamento irregular o operador tem que comunicar imediatamente a equipe técnica de perfuração que deverá executar os procedimentos adequados para solucionar o problema. Geralmente o software do sistema de *mud-logging* permite a programação de alarmes que deverão soar na cabine de *mud-logging*, alertando o operador de *mud-logging* sempre que o valor do parâmetro observado estiver

fora dos limites programados pelo próprio operador para o período em questão.

O número de parâmetros observados pode variar de acordo com características particulares da operação de perfuração. Os parâmetros medidos mais comuns são: profundidade medida do poço, profundidade vertical (TVD), profundidade da broca, taxa de penetração (ROP), peso no gancho (WOH), peso na broca (WOB), torque, rotação da broca (RPM), volume de lama nos tanques, pressão de bombeio de fluido para o poço (SPP), pressão na linha de *Choke*, fluxo de lama, gás total, distribuição da concentração de gás, resistividade do fluido de perfuração, temperatura do fluido de perfuração, dentre outros.

É importante mencionar que alguns dos parâmetros listados acima são realmente medidos por sensores, como o peso no gancho (WOH). Outros parâmetros são calculados a partir dos parâmetros medidos. O peso na broca (WOB), por exemplo, é um parâmetro que é calculado usando-se o valor do WOH e o peso da coluna de perfuração.

Durante a execução da perfuração uma grande quantidade de dados é gerada pelo sistema de *mud-logging*. Devido a dificuldade de se armazenar grandes arquivos, muitas vezes os dados são resumidos para constituírem arquivos menores.

Para evitar problemas com este e ao mesmo tempo reduzir o volume de dados armazenados é interessante ter-se um sistema automatizado que analise os parâmetros de *mud-logging* em cada medição e guarde as informações relevantes segundo esta análise. Paralelamente, os resultados desta análise podem servir como mais um elemento sinalizador para auxiliar o operador, ou mesmo, para alimentar um sistema de tomada de decisão para prevenir anormalidades. O estágio em que se encontra uma operação de perfuração pode ser uma importante informação para a identificação de uma anormalidade. Existem anormalidades que são mais propensas a ocorrer em determinados estágios da operação de perfuração. A partir da análise dos dados do *mud-logging* é possível determinar-se qual é o estágio da perfuração.

3 Estágios Associados à Operação de Perfuração de Poços de Petróleo

A perfuração de um poço de petróleo não é um processo contínuo constituído por uma única operação, mas por uma seqüência de eventos discretos que serão chamados de estágios da operação de perfuração (Bourgoyne *et al.*, 1986). Esta seção apresenta os seis estágios de operações de perfuração usados para implementar um sistema capaz de elaborar automaticamente um registro temporal das operações básicas realizadas durante a perfuração de um poço de petróleo através da interpretação de dados de *mud-logging*, que são mostrados a seguir:

- **Circulando (C):** neste estágio não existe ganho na profundidade do poço. É caracterizada pela

circulação do fluido, pelo alto nível de pressão de bombeio (SPP), um alto nível de peso no gancho (WOH) e uma rotação moderada da coluna de perfuração (RPM). Como não existe corte de rocha, há um nível nulo de peso sobre broca (WOB).

- **Perfurando rotativo (PR):** neste estágio ocorre a perfuração propriamente dita, a broca avança aumentando a profundidade medida do poço. A coluna de perfuração está girando com um alto nível de rotação (RPM) e existe circulação de lama (alto nível de SPP). A coluna de perfuração não está acunhada na mesa rotativa, implicando em um alto nível de peso no gancho (WOH) e nível não nulo de peso sobre broca (WOB).

- **Repassando o poço com o giro da coluna (R):** neste estágio, apesar do alto nível de peso no gancho (WOH), de circulação de lama (SPP) e de rotação da coluna de perfuração (RPM), a broca não avança aumentando a profundidade medida do poço. Nesta situação onde está havendo o repasse de um trecho do poço já perfurado.

- **Perfurando orientado (PO):** a broca avança aumentando a profundidade medida do poço. A diferença aqui é que a coluna de perfuração não está girando (nível nulo de RPM) e a perfuração ocorre devido à atuação da ferramenta de fundo (*downhole motor*). Existe circulação de fluido de perfuração (SPP), um alto nível de peso no gancho (WOH) e nível não nulo de peso na broca (WOB).

- **Repassando o poço sem giro da coluna ou ajustando ferramenta de fundo (AF):** nesta etapa a broca não avança e não aumenta a profundidade medida do poço. Existe circulação de lama (SPP) e um alto nível de peso no gancho (WOH). Esta condição indica que está sendo feito um repasse ou que se está ajustando a ferramenta de fundo.

- **Manobrando (M):** este estágio corresponde à adição ou subtração de uma nova seção à coluna de perfuração. Caracteriza-se pelo acunhamento da coluna de perfuração, causando um baixo nível de peso no gancho (WOH) e com coluna de perfuração permanecendo sem rotação (nível nulo de RPM).

Os seis estágios descritos são estágios básicos da operação de perfuração. Estes estágios foram detalhados considerando-se que as fases são perfuradas com o retorno da lama de perfuração para a superfície. A tecnologia de perfuração considerada foi a perfuração usando motor de fundo e *bent housing*.

4 Modelo do Sistema de Classificação de Operação de Perfuração

O problema de classificação automática de estágios da operação de perfuração consiste em, considerando os parâmetros do *mud-logging*, para cada conjunto de medidas, identificar o estágio da operação segundo as seis classes descritas na seção anterior: 1) circulando (C), 2) manobrando (M), 3) repassando sem giro/ajustando ferramenta (AF), 4) perfurando orienta-

do (PO), 5) repassando (R) e 6) perfurando rotativo (PR).

Embora o registro de *mud-logging* disponibilize uma variada gama de parâmetros, apenas alguns são utilizados para identificar os seis estágios básicos da operação. O especialista, com base na sua experiência e no seu conhecimento técnico, interpreta através de regras próprias as características relevantes e a relação entre alguns desses parâmetros para definir o estágio de perfuração. Os parâmetros e as regras construídas podem variar entre os especialistas.

Nesse trabalho são considerados cinco parâmetros para efetuar a classificação:

- **Profundidade da Broca:** mede qual é a profundidade atual da broca. O cálculo desse parâmetro é efetuado considerando-se a leitura do sensor de altura do gancho. Através da análise desse parâmetro no tempo é possível determinar se o sistema está perfurando. Um aumento efetivo da profundidade é um indicativo de perfuração.

- **Peso no Gancho (WOH):** mede qual é o nível atual de carga no gancho, ou seja, a carga total que o gancho está suportando em determinado instante. A análise desse parâmetro é importante para a identificação das manobras (adição de um novo tubo na coluna de perfuração). Para realizar a manobra a coluna de perfuração é apoiada na plataforma. No momento em que a coluna é apoiada, a carga suportada pelo gancho diminui bastante.

- **Rotação da Coluna (RPM):** mede as rotações realizadas pela coluna de perfuração. A análise desse parâmetro permite analisar se a operação está em estágio rotativo ou não.

- **Pressão de Bombeio (SPP):** mede a pressão na tubulação de bombeio de fluido de perfuração para o poço. Através da análise desse parâmetro é possível identificar se o bombeio de lama para o poço está ou não ocorrendo. Um nível alto de pressão indica que a bomba está ligada.

- **Peso sobre a broca (WOB):** representa a carga efetiva transmitida pela broca à formação rochosa que está sendo perfurada. Ele representa a carga efetiva que está aplicada no contato broca-rocha para auxiliar no processo de corte e esmagamento da rocha. Existe uma relação entre o peso no gancho e o peso sobre a broca. No momento imediatamente anterior ao início do corte, todo o peso da coluna está sendo suportado pelo gancho. Dependendo das necessidades, parte do peso total é transferida para a broca para auxiliar no corte. Observando os parâmetros, nesse caso iremos notar uma redução na carga do gancho e um aumento proporcional no peso sobre a broca.

5 Sistema Imunológico Artificial para Reconhecimento de Padrões

Existem muitos aspectos dos sistemas imunológicos biológicos que servem como metáforas para os siste-

mas computacionais imune-inspirados. A seguir apresenta-se uma breve descrição dos mecanismos naturais imunológicos que norteiam alguns sistemas imunológicos artificiais. Na subseção seguinte descreve-se o algoritmo de classificação imunológico artificial.

5.1 Princípios dos sistemas imunológicos naturais e artificiais

O sistema imunológico protege o corpo contra infecções devido ao ataque de uma agente causador de doença. O sistema imunológico natural oferece duas linhas de defesa: o sistema imune inato e o sistema imune adaptativo. O sistema imune inato consiste de células que podem neutralizar um conjunto pré-definido de agentes infecciosos, ou antígenos, sem requerer uma exposição prévia a eles. O antígeno pode ser um intruso ou parte de células ou moléculas do próprio organismo. Esta parte do sistema imunológico geralmente não é modelada pelos sistemas imunológicos artificiais.

Os animais vertebrados possuem um sistema imune adaptativo que pode aprender a reconhecer, eliminar e recordar de novos antígenos específicos. Isto ocorre como uma forma de seleção natural. A resposta imune adaptativa em sistemas biológicos é baseada em dois tipos de linfócitos (ou células próprias) do corpo: células T e células B.

Quando o sistema imunológico reconhece algum agente causador de doenças o mesmo irá efetuar uma resposta imune para eliminar esse agente do organismo. Uma teoria que explica como ocorre a resposta imunológica é a teoria da seleção clonal (de Castro & Timmis, 2002).

Quando o organismo é invadido por um patógeno, a célula B e a célula T são capazes de reconhecê-lo e de fixar-se a ele para efetuar uma resposta imunológica. Esse antígeno reconhecido ativa a célula B induzindo sua proliferação (expansão clonal) e diferenciação em plasmócitos ou células de memória. A célula B se multiplica dando origem a vários clones que sofrem mutação. A mutação é baseada na afinidade (complementariedade) entre o antígeno e o anticorpo, ou seja, quanto maior for a afinidade entre os dois, menor será a taxa de mutação e vice-versa. Esse processo de mutação acaba fazendo com que alguns clones fiquem mais semelhantes ao antígeno e outros menos. Os clones que forem mais semelhantes ao antígeno são divididos em *plasmócitos*, que são as células que secretam anticorpos para combater a infecção, e *células de memória* que guardam a informação desses antígenos com o propósito de obter uma resposta secundária mais eficiente, ou seja, se ocorrer um novo ataque por este mesmo tipo de antígeno a resposta imunológica será mais eficiente. Esta resposta é conhecida como resposta imunológica secundária.

Estes princípios acima descritos são a base de desenvolvimento de sistemas imunológicos artificiais dedicados à tarefa de reconhecimento de padrões.

5.2 Algoritmo AIRS2 Paralelo

AIRS2 é um algoritmo inspirado no tipo de seleção clonal do sistema imunológico da medula óssea. Ele está relacionado com o desenvolvimento de um conjunto de células de memória que dão uma representação de ambiente aprendido, empregando esquemas de maturação de afinidade e mutação. AIRS2 usa mecanismos de controle da população e o conceito de uma célula B abstrata que representa uma concentração de células B idênticas, e adota o uso de um limiar de afinidade para o mecanismo de aprendizagem.

AIRS2 consiste da descoberta do conjunto de células de memória que podem encapsular os dados de treinamento. Basicamente isto é realizado em um processo de dois estágios, primeiramente envolvendo uma célula de memória candidata e então determinando se esta célula de memória candidata deve ser adicionada ao pool total de células de memória. Este processo pode ser descrito como segue (Watkins & Timmis, 2002):

1. *Inicialização*: Crie uma base aleatória chamada de *pool* de células de memória.
2. *Expansão clonal*. Compare o padrão de treinamento com todas as células de memória da mesma classe e encontre a célula de memória com a melhor afinidade (distância euclidiana) com o padrão de treinamento. Esta célula de memória será referenciada como mc_{match} .
3. *Maturação de afinidade*. Clone e faça a mutação de mc_{match} proporcionalmente à sua afinidade para criar um *pool* de células B abstratas.
4. *Metadinâmica das células B*. Calcule a afinidade de cada célula B com o padrão de treinamento.
5. Aloque recursos para cada célula B de acordo com sua afinidade.
6. Remova as células B mais fracas até que o número de recursos volte ao seu limite inicial.
7. *Ciclo*. Se a afinidade média das células B sobreviventes estiver acima de certo limiar, continue até o passo 8, caso contrário faça a clonagem e a mutação destas células B sobreviventes de acordo com a suas afinidades e volte ao passo 4.
8. *Metadinâmica das células de memória*. Escolha as melhores células B como uma célula de memória candidata (mc_{cand}).
9. Se a afinidade de mc_{cand} para o padrão de treinamento é melhor que a afinidade de mc_{match} , então adicione mc_{cand} ao *pool* de células de memória. Se, além disso, a afinidade entre mc_{cand} e mc_{match} está em certo limiar, então remova mc_{match} do *pool* de células de memória.
10. *Ciclo*. Repita o procedimento a partir do passo 2 até que todos os padrões de treinamento tenham sido apresentados.

Uma vez que a rotina de treinamento é completada, AIRS2 classifica os padrões usando o método dos k vizinhos mais próximos com o conjunto desenvolvido de células de memória.

Em comparação com uma abordagem de mineração de dados, AIRS2 é um procedimento de classificação baseado em clusters.

Algoritmicamente, de acordo com a descrição anterior, a versão paralela de AIRS2 comporta-se da seguinte maneira (Watkins & Timmis, 2004):

- Leia os dados de treinamento no processo principal.
- Divida os dados de treinamento em np processos.
- Execute, em cada processo, os passos de 1 a 9 da versão serial de AIRS2 para o conjunto de dados de treinamento referente a cada processo.
- Reúna as células de memória desenvolvidas em cada processo de volta ao processo principal.
- Una as células de memória reunidas em um único *pool* de células de memória para classificação.

6 Resultados

O sistema de classificação imunológico proposto foi treinado e testado com dados reais de uma base de dados obtida na operação de perfuração de um poço marítimo localizado em um campo petrolífero no Brasil, com uma duração da operação com pouco mais de 63 horas. Nesta base de dados os dados foram registrados minuto a minuto, de modo que cada parâmetro possuía apenas um valor a cada minuto. O número total de amostras registradas foi de 3784 minutos, sendo 75% usadas na fase de treinamento (2838 amostras) do modelo imunológico de classificação e 25% na fase de teste (946 amostras).

Os parâmetros de treinamento do algoritmo AIRS2 Paralelo foram estabelecidos em: número de processos = 5; limiar de afinidade = 0,05; quantidade de recursos = 150; taxa de hipermutação = 2; taxa de clonagem = 10; e taxa de estimulação = 0,9.

O processo de aprendizagem gerou 1563 células de memória: 419 para perfurando rotativo (PR), 239 para repassando o poço com giro da coluna (R), 420 repassando o poço sem giro da coluna ou ajustando ferramenta de fundo (AF), 61 para manobrando (M), 411 para perfurando orientado (PO) e 13 para circulando (C). A porcentagem de redução de dados foi 44,93% e o tempo de construção e treinamento do modelo foi de 3,72 segundos (computador AMD-64).

Os erros estimados durante o processo de aprendizagem dos padrões de operação de perfuração de poços foram 0,0243 para o erro médio absoluto e 0,1559 para o erro médio quadrático.

Os experimentos com o algoritmo AIRS2 Paralelo foram feitos utilizando-se sete (7) vizinhos para o método dos k vizinhos mais próximos.

A seguir são apresentados os resultados da classificação obtidos após o treinamento do sistema de reconhecimento imunológico artificial. Após o processo de aprendizagem, o conjunto de treinamento usado (2838 amostras) foi classificado pelo sistema

de reconhecimento imunológico artificial, obtendo 2631 classificações corretas em comparação com a classificação original feita pelo especialista (Tavares *et al.*, 2005). A taxa de acerto, relação entre o número de classificações corretas e o número total de amostras, foi de 92,7%.

Um conjunto de teste, já classificado pelo especialista, composto por 946 amostras que não haviam sido expostas ao sistema de reconhecimento imunológico artificial durante o processo de aprendizagem, foi submetido ao classificador imunológico conseguindo 883 classificações corretas, com uma taxa de acerto de 93,2%. Estes resultados atestam um bom desempenho do sistema imunológico para esta aplicação, mostrando que ele aprendeu a discriminar as regras que permitem classificar os estágios da operação de perfuração do poço.

Para uma melhor compreensão do mecanismo de aprendizagem do sistema de reconhecimento imunológico artificial, é interessante mapear a distribuição das classificações realizadas por ela, a fim de verificar em quais classes houve uma maior dificuldade de aprendizagem pelo sistema. Uma técnica de visualização bastante útil é a matriz de confusão. A matriz de confusão (ou de contingência) é uma matriz quadrada de números que expressam a quantidade de unidades amostrais, associada a uma dada classe durante o processo de classificação efetuado, e a classe original a que pertencem essas unidades. A matriz de confusão expressa a relação entre os dois tipos de erros associados ao sistema de classificação: os de inclusão e os de omissão. A sua diagonal captura a concordância entre dois mapeamentos. As Tabelas 1 e 2 mostram, respectivamente, a distribuição das classificações realizadas pelo sistema de reconhecimento imunológico artificial para os conjuntos de treinamento e de teste.

Tabela 1 – Matriz de confusão do conjunto de treinamento

		Classe estimada					
		C	M	AF	PO	R	PR
Classificação original	C	9	0	0	3	1	1
	M	0	75	0	0	0	0
	AF	0	0	719	76	0	0
	PO	0	0	14	739	0	0
	R	0	0	0	0	271	72
	PR	0	0	0	0	39	818

Tabela 2 – Matriz de confusão do conjunto de teste

		Classe estimada					
		C	M	AF	PO	R	PR
Classificação original	C	1	0	1	0	0	0
	M	0	22	0	0	0	0
	AF	0	0	242	24	0	0
	PO	0	0	2	251	0	0
	R	0	0	1	0	84	28
	PR	0	0	0	0	6	283

Observando-se as Tabelas 1 e 2, pode-se notar que AIRS2 Paralelo discriminou bem o estágio ‘manobrando’ (M), sem cometer nenhum erro de classificação. Para os estágios ‘repassando o poço sem giro da coluna ou ajustando ferramenta de fundo’ (AF) e ‘perfurando orientado’ (PO) o desempenho do sistema foi satisfatório, com alguns erros de classificação. Porém, AIRS2 apresentou uma dificuldade mais forte de discriminação entre os estágios ‘repassando’ (R) e ‘perfurando rotativo’ (PR), confundindo-os entre si. Isto se deve ao fato dos componentes destes pares de classes não serem linearmente separáveis entre si.

A maior taxa de erro ocorreu na discriminação do estágio ‘circulando’ (C), onde a classificação foi estimada erroneamente para três outras classes, além da classe correta.

Quando o algoritmo AIRS2 Paralelo é usado para o treinamento de conjuntos de dados com uma distribuição de classes desbalanceada, os antígenos das classes com mais amostras podem gerar mais células de memória do que as outras classes com número menor de amostras. Se todas as células de memória são representadas em um alto espaço dimensional, uma classe menor pode ser rodeada por muitas classes majoritárias, de modo que estas classes podem levar os votos de várias células de memória mais próximas e um antígeno pode ocasionar decisões errôneas. Esta é a explicação para o fato do sistema de reconhecimento imunológico artificial não ter conseguido discriminar bem a classe ‘circulando’, produzindo uma baixa taxa de acerto tanto nos dados treinamento quanto no de teste.

Mesmo assim, mesmo para os dados complexos do *mud-logging*, o algoritmo AIRS2 Paralelo é capaz de comportar-se razoavelmente bem como classificador. Todavia, o algoritmo imunológico é ainda inferior ao desempenho do classificador por redes neurais (96,3% para dados de treino e 94,9% para teste) desenvolvido para esta mesma tarefa, com os mesmos conjuntos de dados (Serapião *et al.*, 2005).

7 CONCLUSÕES

O sistema de reconhecimento imunológico artificial mostrou-se satisfatório na classificação dos estágios de operação, atingindo, em média, uma taxa de acerto de 93% dos dados submetidos à discriminação. Este resultado revela um grande potencial de utilização efetiva deste modelo para a automação da monitoração em situações reais durante a perfuração de poços.

A distribuição desequilibrada de dados entre as classes teve um impacto sobre o desempenho do classificador AIRS2 Paralelo, visto que ele pode alcançar uma alta precisão de acerto entre as classes predominantes e uma taxa de acerto mais baixa nas classes com uma quantidade de amostras menor. Mesmo assim, este artigo demonstra que a utilização de abordagens imuno-inspiradas para a classificação de problemas reais multi-classes é viável e factível.

A automação da interpretação de dados de *mud-logging* possibilita a investigação, em tempo real ou *a posteriori*, do desempenho da perfuração do poço, permitindo uma maior avaliação dos requisitos usados durante o processo de perfuração, cuja análise advinda dessa avaliação pode sugerir melhores ajustes dos parâmetros de perfuração para melhorar a eficiência da sonda de perfuração. Outra possibilidade de aplicação deste sistema automático de classificação é utilizar o conhecimento adquirido sobre o tempo gasto em cada estágio de perfuração em um poço de petróleo para aperfeiçoar o processo de perfuração de um novo poço.

Agradecimentos

Os autores gostariam de agradecer o Conselho Nacional de Pesquisa (CNPq) pelo suporte financeiro (Processo CT-Petro/MCT/CNPq 504528/2004-1).

Referências Bibliográficas

- Bourgoyne, A.T., Millheim, K.K., Chenevert, M.E. and Young, F.S., (1986) “*Applied Drilling Engineering*”, SPE Textbook Series, vol. 2, USA, 502p.
- de Castro, L., Timmis, J. (2002) “*Artificial immune systems: A new computational approach*”, Springer, London, UK.
- Serapião, A.B.S., Tavares, R.M., Mendes, J.R.P. e Morooka, C.K. “Classificação automática da operação de perfuração de poços de petróleo através de redes neurais”. Anais do VII Simpósio Brasileiro de Automação Inteligente (SBAI), São Luís-MA, Setembro, 2005.
- Tavares, R.M., Mendes, J.R.P., Morooka, C.K., Plácido, J.C.R. (2005) “Automated Classification System for Petroleum Well Drilling using Mud-Logging Data”. In *Proc. of 18th International Congress of Mechanical Engineer, Offshore & Petroleum and Engineering*, Ouro Preto.
- Watkins A. (2001) *AIRS: A resource limited artificial immune classifier*. M. Sc. Thesis. Department of Computer Science, Mississippi State University.
- Watkins, A., Timmis, J. (2002) “Artificial Immune Recognition System (AIRS): Revisions and Refinements”. In J. Timmis and P.J. Bentley, editors, *Proc. of 1st International Conference on Artificial Immune Systems (ICARIS2002)*, University of Kent at Canterbury, 173-181.
- Watkins, A., Timmis, J. (2004) “Exploiting Parallelism Inherent in AIRS, an Artificial Immune Classifier”. In *Proc. of the 3rd International Conference on Artificial Immune Systems (ICARIS2004)*, Catania, Italy, Lecture Notes in Computer Science (LNCS), v. 3239. Edited by G. Nicosia, V. Cutello, P. Bentley, and J. Timmis, Springer-Verlag, 427-438.