

UM MODELO DE APRENDIZAGEM ADAPTADO AO GERENCIAMENTO TÁTICO DO FLUXO DE TRÁFEGO AÉREO: CASO CINDACTA I

Daniela Pereira Alves

Ministério da Ciência e Tecnologia – MCT
Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT

Bueno Borges de Souza

Li Weigang

Universidade de Brasília - UnB
Departamento de Ciência da Computação – CIC

Antonio Marcio Ferreira Crespo

Primeiro Centro Integrado de Defesa Aérea e Controle do Espaço Aéreo - CINDACTA I

RESUMO

Neste artigo, o modelo de Controle Nível Meta foi modificado e aplicado ao sistema de gerenciamento de tráfego aéreo (caso CINDACTA I: Primeiro Centro Integrado de Defesa Aérea e Controle de Tráfego Aéreo em Brasília) para melhorar o desempenho da comunicação entre os agentes (controladores e supervisores). Durante o processo de controle, a decisão é tomada baseando-se na experiência e no conhecimento adquirido pelos agentes com aprendizagem por reforço. A pesquisa proposta apresenta um Módulo de Avaliação e Apoio à Decisão (MAAD), modelado para um Sistema de gerenciamento de fluxo de tráfego aéreo (ATFM), com a combinação dos algoritmos de aprendizagem por reforço e das técnicas de Controle Nível Meta. Com o sistema desenvolvido, os casos de heurística inicial (HI), epsilon adaptativo (EA) e heurística baseada em desempenho (HD) foram testadas. As análises dos resultados obtidos nas simulações realizadas pelo estudo de caso mostram satisfatoriamente a finalidade da pesquisa.

ABSTRACT

In this paper, the Meta-Level Control model is modified and applied in Air Traffic Flow Management – ATFM (especially in the case of CINDACTA I: First Integrated Center of Air Defense and Air Traffic Control in Brasília) to improve the performance of the communication among the agents (controllers and supervisors). During the control process, the decision made by the agents is based on the experience and there is no decision support system to learn the history knowledge. The research proposed a Model of Evaluation and Decision Support (MEDS) system for ATFM with the combination of the Meta-Level Control approach and reinforcement learning algorithms to assist the controller to get better decision. The paper shows three strategies of reinforcement learning in the simulation, as the cases of initial heuristic (IH), epsilon adaptative (EA) and performance heuristic (PH). The results from simulation and analyses show the satisfactory in the decision process.

1 INTRODUÇÃO

No contexto da crise instalada no setor de transporte aeroviário brasileiro, diversos são os fatores geradores e contribuintes, não cabendo, neste momento, a apresentação em sua totalidade. Contudo, um desses fatores pode estar associado à forma empírica como são utilizadas as técnicas previstas para o controle do fluxo de tráfego, o que fica evidenciada na aplicação intuitiva de medidas restritivas ao fluxo de tráfego aéreo, com o intuito de se evitar a saturação dos setores de controle da Região de Informação de Voo de Brasília - FIR-BS.

A aplicação de medidas restritivas ao fluxo de tráfego é determinada com a finalidade de atender às prescrições regulamentares em vigor, de forma a impedir que os controladores de

vôo trabalhem com carga superior ao estabelecido em legislação, e tem severo impacto sobre todas as atividades operacionais e logísticas aeroportuárias.

A escolha das técnicas e o dimensionamento da sua aplicabilidade, no que tange às medidas restritivas ao fluxo de tráfego, são estabelecidos de forma empírica, fortemente condicionada à experiência dos controladores exercendo as funções de operador e supervisor, ou ainda, mais recentemente, pela consulta direta ao Centro de Gerenciamento de Navegação Aérea (CGNA), que impõem medidas baseadas em estudos estatísticos e em previsões advindas do sistema atualmente usado, o SYNCROMAX. Entretanto, o número de fatores e o número de variáveis envolvidos neste processo tornam a tomada de decisão, empiricamente fundamentada, apenas parcialmente eficaz. As decisões advindas do SYNCROMAX são fundamentadas nos planos de vôos desenvolvidos na fase estratégica e não incluem a real situação que se desenrola no espaço aéreo durante as transições entre os setores aéreos. Ou seja, existe espaço para a otimização do fluxo de tráfego mediante o incremento da eficiência das medidas restritivas aplicadas, o que se dará através da escolha sistematizada da medida mais adequada, considerando-se um grande número de fatores e variáveis envolvidos (capacidade dos setores de controle, dimensão dos setores de controle, demanda por aeroporto, desempenho das aeronaves, necessidade de priorização de aeroportos, necessidade de priorização de rotas, disponibilidade de pessoal, restrições meteorológicas dentre outros), além da inclusão, na previsão dos vôos, dos fatores, circunstâncias e eventos de tempo real.

Diante do que foi apresentado, seria desnecessário mencionar a importância estratégica, econômica e política atribuída ao desenvolvimento de um sistema que possibilite o estabelecimento de medidas restritivas de controle de fluxo mais adequado para o cenário vigente. Este sistema está em fase de projeto e carece do submódulo o qual esse artigo tem como objetivo apresentar.

Este trabalho tem como objetivo apresentar o modelo de um subsistema integrado ao Sistema de Apoio a Decisão aplicado ao Gerenciamento Tático do Fluxo de Tráfego (SISCONFLUX) em fase de projeto pelo Primeiro Centro Integrado de Defesa Aérea e Controle de Tráfego Aéreo (CINDACTA I) e pela Universidade de Brasília (UnB) que auxilie os controladores no gerenciamento das medidas de controle de fluxo de tráfego aéreo.

O artigo apresenta na seção 2 os problemas emergentes apresentados pela análise de requisitos realizada pela equipe de projeto, na seção 3 é apresentada uma visão geral do estado da arte em gerenciamento de tráfego aéreo, a seção 4 apresenta a metodologia empregada na construção do submódulo de apoio à decisão, a seção 5 é apresentada o modelo do subsistema e posiciona este subsistema no sistema completo sendo projetado para apoio à decisão SISCONFLUX, e finalmente na seção 6 apresenta-se um Estudo de Caso obtido através de simulações.

2 PROBLEMAS EMERGENTES

O CINDACTA I dispõe de um conjunto de sistemas capazes de realizar o adequado controle dos movimentos aéreos evoluindo na sua área de responsabilidade. Contudo, não há um sistema específico voltado para o gerenciamento tático e a sincronização do fluxo de tráfego, principalmente na vigência de cenários onde se verifica a degradação dos meios de controle

ou outros fatores que causem modificações significativas no fluxo de tráfego esperado, tais como fenômenos meteorológicos, incidentes e/ou acidentes aeronáuticos, dentre outros.

A degradação dos meios de controle de tráfego, assim como os demais fatores capazes de modificar o fluxo de tráfego esperado, pode levar a saturação dos setores de controle, o que se caracterizam pela permanência simultânea de 14 ou mais aeronaves por setor (Crespo *et al.*, 2007). A saturação de um setor pode ser condicionada por diversos fatores, dentre os quais se destacam: as dimensões do setor, a posição geográfica e o horário do dia. Essa quantidade de aeronaves por setor tem amparo nas normas de definição da carga de trabalhos dos controladores e pode variar para mais ou para menos dependendo dos fatores já citados.

A FIR-BS está subdividida em quatorze setores de controle, agrupados em três regiões: Região Brasília, Região Rio de Janeiro e Região São Paulo. A FIR-BS abarca ainda sete áreas terminais de controle (TMA): Anápolis, Belo Horizonte, Brasília, Cuiabá, Rio de Janeiro, São Paulo e Vitória. Cada setor é controlado por uma posição operacional, onde guarnecem um operador e um assistente. Existe ainda a figura do supervisor, com atuação obrigatória nos momentos em que uma região está sendo controlada por duas ou mais posições operacionais (Rizzi, 2003).

Na eventualidade da saturação de um setor, cabe ao supervisor avaliar a quantidade de tráfegos com previsão de entrada nos setores sob sua supervisão, consultar o CGNA solicitando orientação à medidas restritivas cabíveis e, na manutenção da situação problema, aprovar as medidas restritivas de fluxo de tráfego sugeridas pelo operador. As decisões tomadas pelo supervisor são baseadas nas orientações do CGNA e na sua experiência própria. As medidas restritivas aplicadas decorrem de uma análise dos planos previstos e não incluem dados de tempo real, em situações extremas são necessárias medidas empíricas, desprovidas de suporte computacional de apoio à decisão. Ressalta-se ainda que não é possível, ao supervisor, realizar uma avaliação quantitativa do impacto das medidas adotadas em um determinado setor sobre o fluxo de tráfego nos setores adjacentes. Como consequência, não há um nível adequado de previsibilidade acerca dos efeitos das medidas restritivas adotadas sobre a demanda de fluxo de tráfego da FIR-BS como um todo (Crespo *et al.*, 2007). Como a FIR-BS acomoda cerca de 50% do volume de tráfego aéreo (vôos regulares) no Brasil, o dimensionamento inadequado das medidas aplicadas pelo Centro de Controle de Área de Brasília (ACC-BS) certamente implicará problemas de fluxo de tráfego em todo o Brasil. Estes fatos foram observados durante o processo de levantamento e análise de requisitos do sistema SISCONFLUX.

3 ESTADO DA ARTE PARA ATFM

O gerenciamento de fluxo de tráfego aéreo (*Air Traffic Flow Management – ATFM*) é um problema que envolve sincronização em tempo real (Stoltz e Ky, 2001; Stoltz e Guerreau, 2002). Muitas pesquisas já foram desenvolvidas com o intuito de resolver este problema, porém, como todo problema crítico, as soluções computacionais esbarram na eficiência dos sistemas computacionais, na necessidade de escalabilidade, de alta segurança e sincronização. A maioria dos sistemas já propostos (Gosling, 1987; Karppinen *et al.*, 1991; Schlatter, 1994; Bonzano *et al.*, 1996; Weigang *et al.*, 1997; Targa, 2001; Gonzaga da Silva, 2001; Prevôt, 2002; Nguyen-Duc *et al.*, 2003; Ball *et al.*, 2003; Dell’Olmo e Lulli, 2003; Rizzi, 2003; Mukherjee, 2004; Zhang, 2005 e Liu *et al.*, 2006) apresenta uma arquitetura centralizada, enquanto que (Heymann *et al.*, 2003; Dib, M. V. P. 2004) são soluções com caráter

distribuído. Todas essas soluções apresentam características relevantes na direção da solução eficiente do problema, mas também apresentam problemas relacionados ao desempenho, isto é, o tempo necessário para obtenção da solução e a formação de “gargalos de comunicação”.

A troca de informações aeronáuticas como dados meteorológicos, situações do espaço aéreo, previsões de congestionamento, de forma global, são de crucial importância para melhorar o desempenho e garantir a segurança em todos os diversos segmentos de serviços aeroportuários (IATA, 2003a). Aeronaves e aeroportos são equipados com sofisticados equipamentos que permitem o monitoramento e a rápida transferência de informações estando a aeronave em terra ou em voo (IATA, 2003b).

É através da intuição, associada à experiência, e mais recentemente à consultas ao Centro de Gerenciamento de Navegação Aérea que os controladores, baseando-se nessas informações, decidem ajustes de trajetórias, retardamento de decolagens e pousos, taxa de decolagens em determinado aeroporto etc. É necessário ainda que as trajetórias previstas sejam acompanhadas para determinar quando um par de aeronaves infringirá a distância de separação mínima necessária e, na ocorrência desse problema, qual ação deverá ser tomada para resolvê-lo (Bonzano *et al.*, 1996). As pressões sobre os controladores de voo tendem a aumentar com o aumento das informações que chegam ao seu conhecimento, o “stress” e o desgaste físico e mental dos controladores podem contribuir para algum tipo de erro. Além disso, existe todo um conjunto normativo de regras que procuram limitar a sobrecarga de trabalho imposta aos controladores. As decisões podem conduzir a soluções mais caras dentro de um conjunto de soluções onde poderia existir alguma mais barata, porém a urgência ou o “stress” impede uma análise mais cuidadosa do problema (Bonzano *et al.*, 1996, Crespo *et al.*, 2007).

Neste contexto emerge o problema de espera em solo (*Ground Holding Problem – GHP*), que tem como metodologia de resolução a busca da sincronização entre setores adjacentes, sob a análise do conjunto de pousos e decolagens previstas dentro de uma mesma área de supervisão, de modo que seja garantido o melhor fluxo entre esses setores. Para essa busca, a literatura acerca do GHP sugere a utilização da programação linear inteira (Rizzi, J. A. 2003, Mukherjee, A. 2004 e Ball *et al.* 2003).

Mais recentemente autores vêm sugerindo metodologias baseadas em programação dinâmica para a obtenção de melhores resultados (Dell’Olmo *et al.* 2003 e Zhang *et al.* 2005). Existem ainda modelos distribuídos baseados em técnicas multiagentes (Heymann *et al.* 2003 e Dib, M. V. P. 2004), já citados, que utilizam comunicação por troca de mensagens para negociar um balanceamento global entre os participantes do sistema.

4 METODOLOGIA

4.1 Controle em Nível Meta (CNM)

O controle em nível meta (*Meta-level Control*) é uma camada adicional inserida em sistemas de informação com o objetivo de auxiliar a camada de controle já existente. O CNM procura compensar ou justificar sua existência apontando ao controle caminhos alternativos que melhorem a eficiência do sistema globalmente, garantindo que tarefas de maior utilidade e com prazos de conclusão menores sejam priorizadas, além de procurar melhorar o uso adequado dos recursos computacionais.

O modelo CNM proposto por Lesser e Raja (2005) abrange a construção de um sistema multi-agente onde a camada CNM monitora o controle, por meio de um conjunto de parâmetros, analisando desde as ações que se referem à comunicação entre os agentes até as ações relativas ao conjunto de tarefas escalonadas para execução. O CNM utiliza meta parâmetros para regular as ações de controle do sistema. A análise se baseia nos meta parâmetros para montar uma sequência de estados orientada pelo processo decisório de Markov (Lesser e Raja, 2005), paralelamente é realizada aprendizagem por reforço que tem a função de garantir a aprendizagem de boas ações em detrimento de más ações. O CNM decide quais ações são mais indicadas dadas as configurações atuais do sistema.

4.2 Aprendizagem por Reforço

A aprendizagem por reforço é interessante para ambientes estocásticos. Estes ambientes sofrem alterações e as ações que são mais indicadas em um determinado instante podem ser inadequadas em instantes posteriores devido à natureza estocástica do ambiente. O aprendizado é decorrência da experiência adquirida por meio de sua interação com o ambiente (Russel *et al.*, 2004).

O agente interage com o ambiente diretamente para obter informações, e estas são processadas através de um algoritmo apropriado, objetivando executar a ação que maximize a obtenção de resultados positivos. O agente atua no ambiente e recebe um valor denominado reforço que é retornado como resposta diante da ação tomada (Sutton *et al.* 1998). Na aprendizagem por reforço o agente pode *usufruir* o conhecimento que já adquiriu com a experiência ou *explorar* novas ações podendo estas retornar melhores ou piores resultados. O objetivo é, geralmente, otimizar algum comportamento dentro do ambiente.

Dois algoritmos ocupam posições de destaque dentre os algoritmos de aprendizagem por reforço: Q-learning e SARSA. O agente de aprendizagem que emprega o algoritmo Q-learning (QL) (Watkins *et al.* 1992) busca maximizar a recompensa obtida pelo agente usando a ação que anteriormente lhe deu a melhor resposta para determinado estado ou explorar ações de maneira aleatória de forma a pesquisar as ações de melhor retorno para o estado em questão. O agente de aprendizagem que emprega o algoritmo SARSA (Sutton *et al.*, 1998) sorteia uma ação a ser executada baseando-se na distribuição de probabilidade das ações tomadas pelo agente. Desta forma, existe uma chance maior de que a ação mais tomada, até então, seja tomada novamente.

4.3 Aprendizagem por Reforço (AR) acelerada por Heurística

A aprendizagem por reforço acelerada por heurística combina as vantagens de AR com o desempenho alcançado com uso de heurísticas. AR aprende através da experiência o que é bom em ambientes estocásticos. Nesses ambientes, a melhor ação depende do estado observado em determinado momento e não há previsão, *a priori*, de quais são as melhores ações a serem tomadas antes que o reforço sobre a ação seja obtido. Fatores aleatórios interferem nos resultados alcançados à medida que as interações do algoritmo são realizadas. Heurística consiste numa série de conhecimentos que proporcionam uma rápida solução para algum problema ou dificuldade, com o menor gasto de energia ou esforço (Wikipedia, 2007). As heurísticas, no caso dos algoritmos de aprendizagem por reforço, têm a função de tornar o processo de aprendizagem mais rápido (Bianchi, R. A. C., 2005).

4.4 Metodologia Proposta

Além dos algoritmos Q-learning e SARSA, algumas melhorias foram propostas com intuito de se acelerar o aprendizado: *heurística inicial*, *epsilon adaptativo* e *heurística baseada em desempenho*. Elas são explicadas nos parágrafos seguintes.

A primeira estratégia proposta, *heurística inicial* (HI), combina as sugestões presentes em Lesser e Raja (2005). Para tanto, uma estrutura de dados armazena a ação considerada mais indicada a ser tomada pelo controle conforme o estado atual do ambiente. As regras conduzirão o sistema ao armazenamento de estados e, respectivas, ações que conduzam a um melhor equilíbrio na distribuição das ocupações dos setores com o mínimo de atraso possível. No contexto do CINDACTA I isto será refletido no aumento da taxa de decolagem dentro dos terminais.

A segunda estratégia proposta, *epsilon adaptativo* (EA), tem caráter técnico relaciona-se com o aumento ou diminuição da taxa de exploração do algoritmo. Procura adequar o “*trade-off*” existente em algoritmos de aprendizagem por reforço que é escolher entre usufruir o conhecimento adquirido até o momento ou explorar novas ações com o intuito de atingir melhores resultados. Se bons resultados estão sendo obtidos, aumenta-se a probabilidade de usufruir, por outro lado, se resultados ruins estão sendo obtidos, diminui-se a probabilidade de usufruir, aumentando-se, por consequência, a taxa de exploração. O Desempenho do sistema é medido com base na ocupação dos setores e na taxa de decolagem nas terminais.

A terceira estratégia proposta é a inserção de uma heurística baseada em desempenho (HD) do sistema. Esta técnica fundamentada no trabalho de Bianchi (2005). Um valor de desempenho ideal é considerado e tomado como referência, o valor de desempenho atual é avaliado e comparado ao ideal. A política de exploração do algoritmo de aprendizagem fica definida segundo o critério de desempenho. A função heurística definida por Bianchi é então representada pela distância entre o desempenho atual e o desempenho idealmente desejado. O valor obtido com essa diferença interfere no usufruto do agente de aprendizagem. O objetivo desta heurística é minimizar a distância entre o desempenho atual e o desempenho tomado como referência. Em todas as melhorias propostas (EA e HD), a estratégia *heurística inicial* HI é aplicada no início do processo.

5 MODELO DO SUBSISTEMA

O sistema completo será capaz de sugerir ao supervisor de uma região de controle de tráfego, as medidas restritivas de controle de fluxo de tráfego mais efetivas diante de determinado cenário vigente. Tais medidas representarão a solução mais adequada para a manutenção da melhor condição de fluxo de tráfego na FIR-BS como um todo.

O SISCONFLUX também deverá ser capaz de sugerir medidas restritivas de fluxo assumindo condições de fluxo de tráfego forçadas, segundo as quais determinados vôos, rotas ou aeroportos poderão ser priorizados em detrimento de outros. Tais priorizações serão estabelecidas conforme as orientações emanadas dos órgãos responsáveis pelo gerenciamento do fluxo de tráfego no espaço aéreo brasileiro.

Fazendo parte deste sistema está o Módulo de Avaliação e Apoio à Decisão (MAAD) que funcionará como um módulo de análise e aprendizagem. Ele terá como função principal sugerir ao Supervisor a medida restritiva de fluxo de tráfego mais adequada ao cenário

projetado, de forma que os setores de controle de tráfego não atinjam a condição de congestionamento ou saturação. A medida sugerida, portanto, será definida dentre as possibilidades levantadas pelo Módulo de Balanceamento de Fluxo.

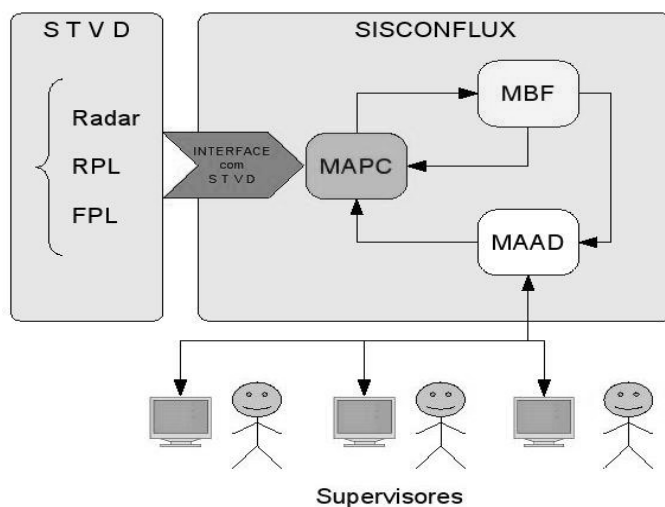
A definição da medida restritiva mais adequada poderá ser condicionada a aplicação de parâmetros específicos que serão definidos pelo Supervisor, tais como a necessidade de priorização de determinados vôos ou de todos os vôos de determinado aeroporto. Para tanto, o subsistema deverá ser alimentado com estes parâmetros, que serão enviados a outros módulos integrantes do sistema para a projeção de novo cenário, projeção de um novo balanceamento de fluxo e a execução de uma nova análise.

O Módulo de Avaliação e Apoio à Decisão fará a interface entre os agentes humanos e os outros módulos do sistema. Fundamentalmente ele tem a função de captar a experiência dos supervisores e controladores de tráfego aéreo, incluindo estas variantes ao conjunto de resultados analisados pelos módulos anteriores, além de funcionar como origem de dados para o sistema como um todo.

Uma vez que as decisões sejam aceitas pelos supervisores, as mesmas serão armazenadas e passarão a fazer parte do sistema como uma decisão aceita para o dado cenário. Dentre as tarefas de responsabilidade desse módulo pode-se citar:

1. Apresentação dos resultados aos supervisores;
2. Coleta de dados para fins de aprendizagem;
3. Armazenamento do conjunto de pares (estado; ação);
4. “Log” de eventos como ações de operadores;
5. Atividades que sejam consideradas úteis administrativamente (item em discussão com o comando do CINDACTA I).

Para apoiar a decisão, a técnica de Controle Nível Meta (*Meta-level Control*) será modificada e aplicada ao sistema para melhorar o desempenho da comunicação entre os agentes. Durante o processo de controle, a decisão é tomada baseando-se na experiência e no conhecimento adquirido pelos agentes com aprendizagem por reforço.



Onde:

SISCONFLUX: Sistema de aplicação e gerenciamento das medidas de controle de fluxo de tráfego aéreo.

MAPC: Módulo de Acompanhamento e Previsão de Cenário.

MBF: Módulo de Balanceamento de Fluxo.

MAAD: Módulo de Avaliação e Apoio à Decisão

STVD: Sistema de Tratamento e Visualização de Dados.

RPL: Planos de vôos repetitivos.

FPL: Planos de vôos eventuais.

Figura 1: SISCONFLUX

O módulo de aprendizagem MAAD inserido na arquitetura do SISCONFLUX (Figura 1) garantirá a capacidade de adaptação em situações estocásticas, ou seja, adaptação às alterações do ambiente que exijam uma mudança no procedimento decisório atual, pelo uso de aprendizagem por reforço acelerado por heurísticas. O módulo de aprendizagem conta com um algoritmo de aprendizagem por reforço que pode ser SARSA ou Q-learning.

A figura 2 mostra um esquema de funcionamento do subsistema e seu relacionamento com os outros módulos do SISCONFLUX.

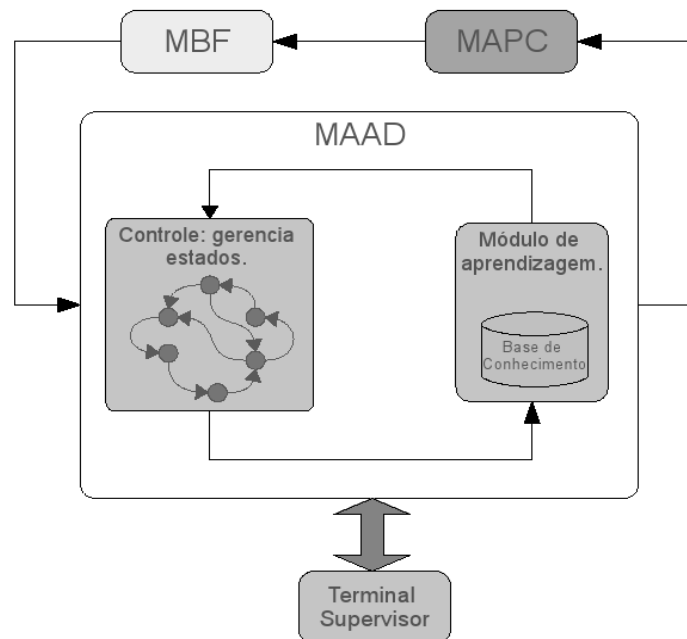


Figura 2: Esquema de relacionamento e arquitetura interna do MAAD

6 ESTUDOS DE CASO

O estudo de caso apresentado neste artigo avalia a velocidade de aprendizagem do módulo de aprendizagem dentro de um ambiente estocástico e também, o conjunto e a qualidade das decisões tomadas. O problema em análise lida com dois aspectos que se contrapõem. De um lado existe a necessidade de se tomar decisões rápidas e, por outro lado a qualidade das decisões que são tomadas. Acontece que para tomar boas decisões é exigido um tempo de raciocínio sobre as opções disponíveis e nem sempre existe o tempo necessário.

6.1 Simulação

O agente de aprendizagem foi submetido a um relacionamento com três outros agentes que, para ele, são fontes geradoras de situações diversas. Podem-se abstrair essas fontes geradoras como sendo as ações do módulo de balanceamento ou as decisões tomadas pelos supervisores no intuito de resolver os problemas que se apresentam. Em geral, não existem diferenças entre a lógica de cada fonte geradora. A intenção é que elas representem as adversidades que ocorrem no dia-a-dia do Centro de Controle, juntamente com as ações rotineiras que se repetem de tempos em tempos e que conduzem a uma solução interposta pelo supervisor. As adversidades enfrentadas pelos aeroportos são simuladas através do aumento ou diminuição do número de problemas que se apresentam ao supervisor. O contexto em análise está sujeito às intempéries relativas a comunicação. Assim, para representar um processo de troca de mensagens intensa, as mensagens, representantes das situações a serem resolvidas são geradas

com espaçamento menor entre uma e outra mensagem. De outra forma, um processo de troca de mensagens de intensidade leve é representado com espaçamento maior entre as mensagens. Para avaliar a qualidade da aprendizagem é estabelecida uma nota ao agente que define, em determinado instante, o quão boa está sua aprendizagem em relação a um padrão considerado como ideal (veja figura 3). Na figura 3, o eixo das ordenadas indica uma nota de zero a cem, atribuída à ação tomada pelo agente em determinada circunstancia. Quanto maior essa nota, maior foi a proximidade com a ação ideal para o problema. Uma nota é gerada a partir da ação tomada em relação ao que é sugerido pela heurística inicial e também considerando o grau de tensão com que as mensagens são produzidas. Neste segundo caso, avaliações da qualidade do aprendizado, tanto a quantidade de períodos quanto o intervalo de geração entre as mensagens foram variados através de oito simulações. Como a intenção não é comparar os algoritmos Q-learning e SARSA, já exaustivamente estudados pela comunidade acadêmica, optou-se por testar as melhorias propostas por esse artigo, no intuito de estabelecer que melhoramentos devam ser usados.

A heurística inicial (HI) foi utilizada em todas as simulações, porque se verificou que em sistemas críticos, como é o caso, não é admitido que o agente parta de conhecimento zero, ou seja, ele deve ser iniciado com as ações mais indicadas previamente. A figura 3 mostra os testes realizados com epsilon adaptativo (QL+HI+EA), heurística baseada em desempenho (QL+HI+HD) e a combinação de todas (QL+HI+EA+HD) sobre o algoritmo Q-learning (QL). A figura 4 mostra os mesmos testes sobre o algoritmo SARSA.

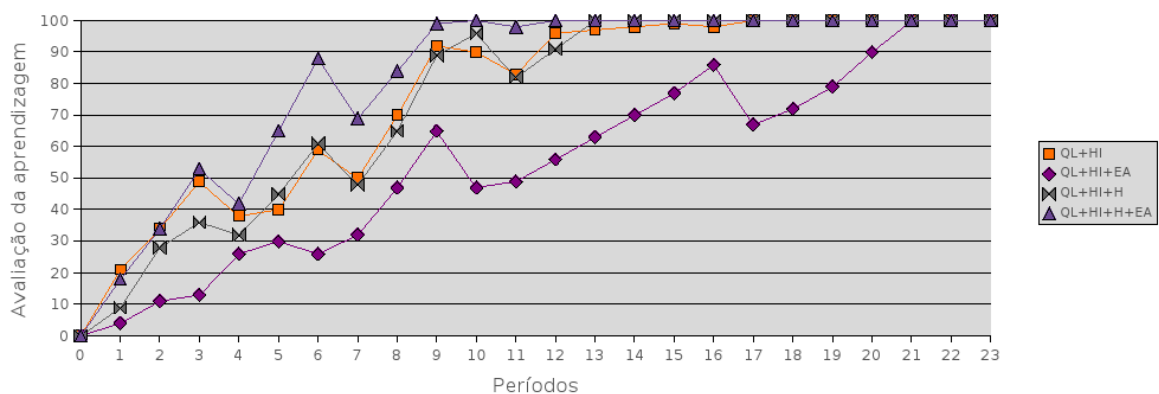


Figura 3: Melhorias propostas aplicadas ao Q-learning.

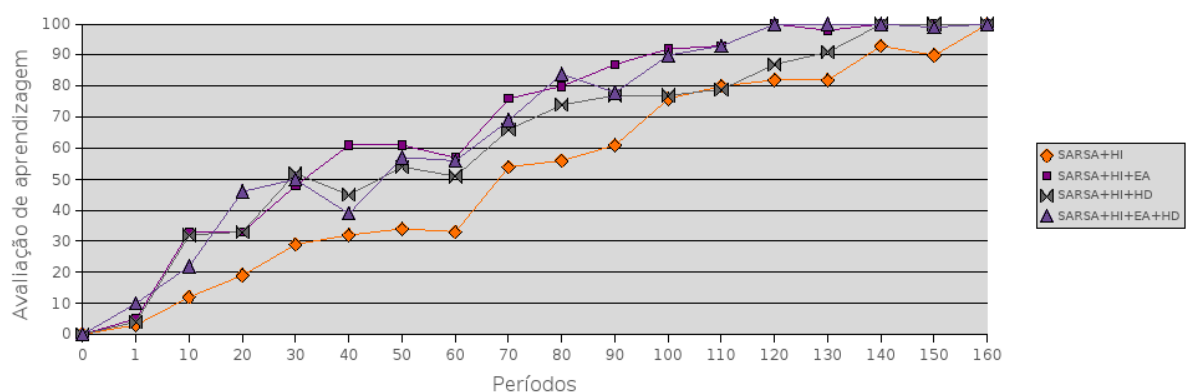


Figura 4: Melhorias propostas aplicadas ao SARSA.

Verificou-se que apenas a melhoria de adaptar o fator epsilon no algoritmo Q-learning não produz bons resultados, pois diminui a velocidade de aprendizagem do algoritmo (note a curva de aprendizagem). Observa-se também que a Heurística baseada em desempenho mantém a basicamente a mesma velocidade. Um pequeno aumento de velocidade na aprendizagem ocorreu com a combinação da Heurística baseada em desempenho com a adaptação do parâmetro epsilon. Quanto ao algoritmo SARSA verificou-se que as alterações produzem uma pequena aceleração na aprendizagem que em condições críticas podem ser de grande importância, principalmente em ambientes estocásticos onde as mudanças ocorrem com bastante frequência. Como é notório pela comunidade acadêmica, o algoritmo SARSA tem mais velocidade de processamento, porém precisa de mais tempo para atingir um bom nível de aprendizagem (figura 4), enquanto que o algoritmo Q-Learning (figura 3) apresenta um tempo maior de processamento, mas consegue ótimos resultados de aprendizagem em poucos ciclos. No sistema real tem-se a combinação dos dois algoritmos como um ponto chave para a obtenção de rapidez tanto na aprendizagem como no processamento.

6 CONSIDERAÇÕES FINAIS

A principal contribuição deste trabalho está na apresentação de um modelo de gerenciamento e apoio à decisão que emprega, ao mesmo tempo, estratégias de IA, controle em nível meta e aprendizagem por reforço. Ao considerar a aprendizagem por reforço, houve contribuições com a apresentação de adaptações aos algoritmos Q-learning e SARSA considerando a importância de avaliar a qualidade da decisão tomada e também, a velocidade com que uma decisão é tomada. Outra contribuição importante ao contexto considerado foi a hierarquização de atividades permitindo que situações mais críticas possam ser tratadas de maneira mais rápida e efetiva.

Como trabalhos futuros, sugerem-se a integração desses procedimentos com os procedimentos previstos nos demais módulos integrantes do sistema SISCONFLUX no intuito de auxiliar as atividades dos controladores e supervisores no processo de decisão. Sugere-se, ainda, o aprimoramento da heurística baseada em desempenho para refletir com maior precisão os critérios de desempenho relativos à FIR-BS.

O sistema SISCONFLUX constitui objeto de um projeto estratégico de cooperação entre o Comando da Aeronáutica e a Universidade de Brasília, com vistas ao desenvolvimento de um sistema de extrema importância no contexto operacional dos órgãos de controle de tráfego e que elidirá uma importante lacuna dentro da concepção nacional de gerenciamento do tráfego aéreo. Neste contexto, o uso de aprendizagem por reforço para acúmulo de experiência do supervisor se mostra fundamental.

A equipe de projeto apresenta um caráter multidisciplinar e é integrada por pesquisadores da UnB, do CINDACTA I e da Comissão para Implantação do Sistema de Controle do Espaço Aéreo (CISCEA). O Projeto, com prazo de conclusão estimado em três anos contados a partir de 2007, já conta com importantes aportes financeiros e de equipamentos, além de manter um intenso intercâmbio com o Centro de Gerenciamento da Navegação Aérea (CGNA), órgão do DECEA responsável pela concepção e execução do gerenciamento de tráfego aéreo no Brasil.

O potencial das idéias apresentadas neste artigo não se limita ao controle de fluxo de tráfego aéreo. Pode-se estender essa solução para muitos problemas em inteligência artificial cujo cunho seja o aprendizado por meio da experiência.

REFERÊNCIAS BIBLIOGRÁFICAS

- Ball, M. O.; Hoffman, R.; Odoni, A. e Rifkin, R (2003), *A stochastic integer program with dual network structure and its application to the ground holding problem*, Institute for Operations Research and the Management Sciences (INFORMS) 51 (2003), 167–171, ISSN:0030-364X.
- Bianchi, R. A. C. (2005). *Uso de Heurísticas para a Aceleração do Aprendizado por Reforço*. XXV Congresso da Sociedade Brasileira de Computação. São Leopoldo.
- Bonzano, Andrea; Cunningham, Pádraig; Meckiff, Colin (1996) *ISAC - A CBR System for Decision Support in Air Traffic Control*. August (In I. Smyth & B. Faltings (Eds.), *Advances in Case-Based Reasoning*. Amsterdam: Springer Verlag).
- Crespo, M. F. C.; Aquino, C. V.; Souza, B. B.; Weigang, L.; Melo, A. C. M. A.; Alves, D. P. (2007) , *Sistema Distribuído de Apoio a Decisão aplicado ao Gerenciamento Tático do Fluxo de Tráfego: Caso CINDACTA I*, VI Simpósio de Transporte Aéreo - SITRAER, Maringá PR, volume 1, 317-327.
- Dell’Olmo, P e Lulli, G (2003), *A dynamic programming approach for the airport capacity allocation problem*, IMA Journal of Management Mathematics 14 (2003), 235–249.
- Dib, M. V. P. (2004) *Sistema Multiagentes para Sincronização e Gerenciamento de Fluxo de Tráfego Aéreo em Tempo Real*; Dissertação de Mestrado, Universidade de Brasília.
- Gonzaga da Silva, A. M. (2001) *Sistema de simulação acelerado para análise de fluxo de tráfego aéreo*. Master’s thesis, Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, SP, Brasil.
- Gosling, G. D. (1987) *Application of Expert System in Air Traffic Control*, Transp. Eng. V113, No 2, pp. 139-155.
- Heymann, M., Meyer, G. e Resmerita, S (2003), *A framework for conflict resolution in air traffic management*, Proceedings, 42nd IEEE Conference on Decision and Control (Maui, Hawaii), vol. 2, December 2003, Digital Object Identifier 10.1109/CDC.2003.1272914, pp. 2035–2040.
- IATA (2003a) *Aeronautical Information and Meteorological Data Format*. Eleventh Air Conference. Montreal, 22 de setembro de 2003.
- IATA (2003b) *IATA Position on the Global Air Traffic Management (ATM) Operational Concept and the Need for an ATM Global Implementation Roadmap*. Eleventh Air Conference. Montreal, 2003.
- Karppinen, N.; Lucas, A.; Ljungberg, M.; Repussseau, P. (1991) *Artificial Intelligence in Air Traffic Flow Management*. Anais International Aerospace Conference. Melbourne, Australia.
- Lesser, V. e Raja, A. (2005) *Automated meta-level control reasoning in complex agents*. IEEE.
- Mukherjee, A (2004), *Dynamic stochastic optimization models for air traffic flow management*, Ph.D. thesis, University of California, Berkeley, 2004, UCB-ITS-DS-2004-3.
- Nguyen-Duc, M.; Briot, J. P.; Drogoul, A. e Duong, V. (2003) *An application of Multi-Agent Coordination Techniques in Air Traffic Management*, in the Proceedings of the IEEE/WIC International Conference on Intelligent Agent Technology (IAT’03).
- Prevôt, T. (2002) *Exploring the Many Perspectives of Distributed Air Traffic Management: The Multi Aircraft Control System MACS*. In S. Chetty, J. Hansman, and G. Boy (Eds.), *Proceedings of (HCI-Aero 2002)*, AAAI Press, Menlo Park, CA, pp. 149-154.
- Rizzi, J. A. (2003) *Um modelo matemático de auxílio para o problema de controle do tráfego aéreo*. Master’s thesis, Instituto Tecnológico de Aeronáutica - ITA, São José dos Campos, SP, Brasil.
- Russel, S. e Norvig, P. (2004) *Inteligência Artificial*. Editora Campus.
- Schlatter, U. R. (1994) *Real-Time Knowledge-based Support for Air Traffic Management*. IEEE Intelligent Systems. Junho de 1994 Vol. 9 n° 3 pp 21-24.
- Stoltz, S. e KY, P. (2001) *Reducing Traffic Bunching More Flexible Air Traffic Flow Management*, 4th USA/Europe ATM R&D Seminar, New_Mexico.
- Stoltz, S. e Guerreau, R. (2002) *EUROCONTROL: Future ATFM Measures (FAM) operational Concept*, EEC Note No. 13/02.
- Sutton, R. e Barto (1998) *A Reinforcement Learning – An Introduction*. A Bradford Book, Cambridge, Massachusetts.
- Targa, D. (2001) *Uma ferramenta automatizada no auxílio à alocação de slots para o problema de gerenciamento de fluxo de tráfego aéreo brasileiro*. Master’s thesis, Instituto Tecnológico de Aeronáutica - ITA, São José dos Campos, SP, Brasil.
- Watkins, Christopher J.C.H.; Dayan, Peter (1992) *Technical Note: Q-Learning*, *Machine Learning* 8:279-292.
- Weigang, L.; Alves, C. J. P. e Omar, N. (1997) *An expert system for Air Traffic Flow Management*. *Journal of Advanced Transportation*, Vol. 31, No. 3, 1997, pp. 343-361.
- Wikipedia (2007) Acessado em 03/07/2007 pelo link <http://en.wikipedia.org/wiki/Heuristic>
- Zhang, Z; Gao, W e Wang L. (2005), *Short-term flow management based on dynamic flow programming network*, *Journal of the Eastern Asia Society for Transportation Studies* 6 (2005), 640–647.